COOPERATIVE PROTOCOLS IN DENSE WIRELESS NETWORKS FOR BROADCAST AND CONSENSUS

by Shrut Kirti

COOPERATIVE PROTOCOLS IN DENSE WIRELESS NETWORKS FOR BROADCAST AND CONSENSUS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Shrut Kirti

August 2010

COOPERATIVE PROTOCOLS IN DENSE WIRELESS NETWORKS FOR BROADCAST AND CONSENSUS

Shrut Kirti, Ph.D.

Cornell University 2010

Network congestion is a communication bottleneck in large wireless networks which use packet-switched communications, impeding their performance in critical applications like network broadcast and in-network data processing. This forms the motivation for this thesis. We propose low complexity physical layer communication protocols that use cooperative transmission to ameliorate this problem. We show that cooperation at the physical layer can significantly improve performance with relatively small coordination overheads by exploiting the broadcast nature of the wireless channel.

In the first part of the thesis, we address congestion in broadcast and propose decode and forward protocols that can accommodate multiple users broadcasting their content simultaneously. These protocols incorporate state of the art techniques like power control, successive interference cancellation, use of side-information for decoding, and interference alignment. By considering the two user linear network in depth we obtain the necessary conditions for the successful broadcast of the content of multiple users. It is shown that these physical layer cooperative protocols achieve lower broadcast latency than packet-switched flooding protocols and are suitable for broadcast in environments with fading. Additionally, we examine the performance of single user broadcast in the presence of decoding errors and show that the errors in its information flow need not be catastrophic. Supporting numerical results are included.

In the second part of the thesis, we employ a similar approach and propose an efficient physical layer architecture for average consensus gossiping algorithms which perform in-network computation. This architecture relies on structured codes which combine channel and source coding. These codes result in consensus updates that are data driven, where transmissions are scheduled based on the states of the nodes rather than their index. Through this simple strategy we show that in spite of bandwidth and power limitations, nodes in an increasingly dense network can converge to the average with bounded delay and precision. Simulations show that this strategy outperforms packet-switched protocols even for moderately sized networks.

BIOGRAPHICAL SKETCH

Shrut Kirti was born in New Delhi, India on December 23, 1981. He received his B.S. degree in Electrical Engineering from The University of Texas at Austin, Texas, USA, in 2005. Since then, he is has been working toward his Ph.D. degree at Cornell University, Ithaca, NY. He received his M.S. degree, also from Cornell University, in 2008. His general research interests lie in distributed algorithms for sensor networks, cooperative relay networks, and physical layer protocols for large networks.

To my family.

ACKNOWLEDGEMENTS

It is my sincere pleasure to thank my advisor, Prof. Anna Scaglione, for her guidance, ideas, and encouragement. Her infectious enthusiasm and skill for solving problems is inspiring, her drive for constant improvement in one's work humbling. I'll always value the lessons I have learned from her.

I'd like to thank my thesis committee members Prof. Lang Tong and Prof. Kevin Tang for useful suggestions on my thesis. I'm grateful to all my other instructors from whom I've been fortunate to learn so much. I'd also like to thank Prof. Robert Thomas and Prof. Bhaskar Krishnamachari who took the time to provide guidance when I sought it.

I will remember all the memorable times I have spent with my group mates and friends - Matthew Sharp, Ercan Yildiz, Roberto Pagliari, and Ramy Tannious. Their friendship made my studies much more enjoyable. I'd like to thank Sasikanth for the many interesting meals, discussions, and arguments, and folks at the Cornell badminton club, for becoming great friends and improving my game.

My parents provided constant support through thick and thin, this has been invaluable. It was a great convenience that my brother Divya joined me at Cornell for his own studies. That we both completed are studies together is even better. I give my deepest thanks to my family, my parents and my younger brother, without whose love this thesis would not be possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Spurred by innovations in low cost computing, wireless communication, and other hardware, it is quickly becoming possible to deploy networks on a large scale, *e.g.* [46, 36]. The deployment of such large scale networks is only expected to increase further in the coming years. In order to fully exploit the real-time capabilities of these large networks, however, it is imperative to study techniques for, first, quickly moving information around within the network and second, efficiently accessing and processing the information collected in disparate network locations. We address both these important issues in this thesis.

One of the fundamental communication primitives in any networking protocol is broadcast. It is used to disseminate control messages, advertise or discover network services, establish unicast routes, or multicast in hostile environments when there are no other communication alternatives. Moreover, in wireless networks with mobile and/or resource constrained nodes, like ad hoc or sensor networks, and in large scale networks, the cost of routing [28] can be prohibitive. Having efficient broadcasting protocols is therefore essential. Additionally, as wireless devices emerge as media platforms, broadcasting content with delay guarantees over multiple cells may enable peer to peer distribution services to run on the air (Fig. 1.1). Unfortunately, traditional packet-switched network flooding protocols fall short of providing low latency methods for delivering packets because of the *broadcast storm* problem [59]. Although this is a compelling problem to solve, few physical layer or coding solutions have been advocated for scalable network

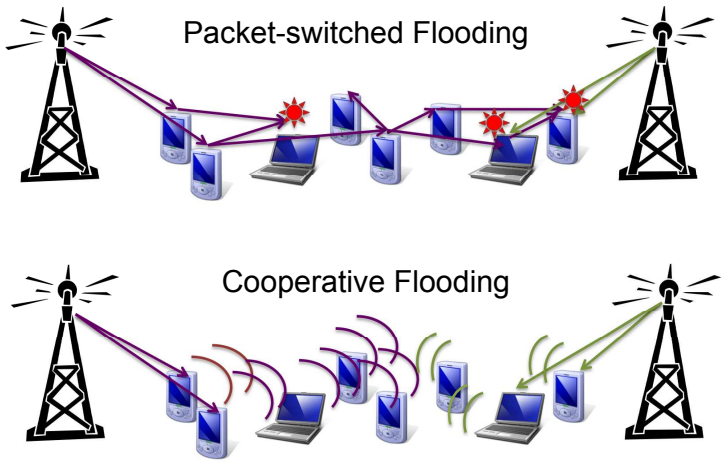


Figure 1.1: Packet-switching protocols suffer from the broadcast storm problem. The physical layer cooperative protocols we propose do not.

flooding. We propose physical layer cooperative decode and forward protocols for multiplexing the broadcast information of several sources in the network which we show outperform non-cooperative packet-switched flooding protocols.

As regards accessing and processing network data, network architectures which utilize a fusion center are naturally constrained - the fusion center is a communication bottleneck and a single point of failure. Gossiping protocols [66, 41, 11, 63] on the other hand, offer an alternative by relying on repeated near-neighbor communications across the network. Such protocols are attractive in wireless sensor network applications for a number of reasons. For example, they alleviate the complexity of the network layer. The low control overhead of local communications is well suited to the low operating power restriction of sensors. In addition, gossiping protocols are robust to node failures and possible network topology changes since they can adapt to such modifications. But gossiping protocols are no panacea. While they are not constrained by a single point of failure, they are susceptible to *communication* bottlenecks. These bottlenecks can be described in different terms depending on the abstraction chosen to analyze the network. Possible bottleneck

measures are: a limit on the number of messages exchanged, packet drop rate, the number of bits per message used, or delay.

In this dissertation we focus on a particular type of gossiping - *average consensus* - which is aimed at computing the average of the states of all nodes distributedly [11, 63, 21]. In the case of average consensus, in order to achieve a certain precision in the consensus estimate in a given time, each node should successfully receive a certain average number of updates from its neighbors. The rate at which a node can receive these updates is a communication bottleneck. As the network size and the number of neighbors increases, the main source of this bottleneck is congestion since greater traffic in the channel is associated with increased packet drop and delay, especially in wireless networks with interference. Network congestion can be mitigated by increasing the available spectrum, but unfortunately, increasing transmission power yields no benefit. Therefore, even if we neglect the problem of transmitting under a rate constraint, computing the average of states in an infinitely dense network over a band-limited wireless channel using standard scheduling is impossible due to congestion. We show that by rethinking conventional assumptions on the communication architecture it is possible to prove this conclusion invalid. We show that this limitation disappears if we combine the computation with source and channel coding by using a cooperative data driven update strategy.

1.2 Cooperative Multicasting

The first part of this thesis considers broadcasting the content of multiple (M) sources simultaneously in a large network using physical layer decode and forward

cooperative protocols. Briefly, in the first time-slot, M users broadcast their messages and all nodes attempt to decode them using either multi-user or single-user receiver techniques. In the next slot, all nodes that decoded a particular message correctly, forward that message. This process repeats. When multiple nodes are be able to decode and therefore forward the same message, symbol-synchrony naturally leads to the emergence of cooperative transmission because we do not try to avoid collisions of signals carrying the same message at the physical layer. On the other hand, as shown in Fig. 1.1, packet-switched networks will incur ever greater cascading collisions if multiple sources are broadcast simultaneously. It is natural to wonder if cooperative broadcasting will perform better: Will cooperative flows collide and stop or will the flows intersect and continue to propagate due to opportunistic capture effects[1] that allow the flows to persist?

This thesis shows that the reinterpretation of interference in this manner coupled with opportunistic gains improves broadcast latency by avoiding channel contention. Further gains are realized by the using of multi-user techniques like Successive Interference Cancellation (SIC), which allow efficient information multiplexing. The protocols are fully distributed and do not incur any overheads like routing or relay selection. Moreover, they are scalable to networks with high node densities or mobility since individual node addressing is not required. The protocols are fair - each node utilizes the same amount of energy for relaying. The caveat is that we assume symbol-level network synchronization (see Section 2.1 for further discussion).

While the strategies we propose are general, in our analysis we focus on the case where two source nodes are placed at opposite ends of a linear network with

[1]In dense networks with fading, the protocols benefit from spatial diversity where "lucky" nodes with good fades are able to decode messages they would ordinarily fail to, leading to opportunistic gains.

an asymptotically large number of nodes. We draw insight from these analytical results and provide necessary conditions for the case of M sources.

We also address the issue of what happens when errors are introduced in the cooperative relaying process, although for tractability we limit our studies in this scenario to a single source. Not many authors have ventured into the analysis of the error propagation phenomenon in large cooperative networks. Several authors avoided the issue by invoking information theoretic limits [45, 48]. Others only considered small scale networks [15, 27]. In all cases the expressions are challenging to interpret. On the other hand, it is quite easy to obtain scaling laws for non-cooperative multi-hop systems (see Section 3.2.3) and to show that the probability of error degrades steadily with the number of hops and, thus, with the distance covered by the message. The obvious question to ask is if this is also the case in cooperative broadcast and if cooperative broadcast has worse error propagation. Intuition seems to be of little help because in cooperative broadcast erroneous signals are mixed with the correct ones at the physical layer.

To answer this question, we consider two cooperative broadcast protocols similar to the ones considered in [51, 77] (see Section 3.1) for the case of uncoded binary transmission. In both protocols contiguous groups of nodes relay a BPSK modulated message from a source at one edge of the network to all other nodes in steps. One scheme allocates orthogonal channels for all the nodes that cooperate while the second scheme forces all nodes in each group to share the same channel. Using techniques first introduced in [51] for cooperative networks, we analyze the error propagation by using the limit distribution of the received signal in the asymptotic regime, *i.e.,* the number of nodes that cooperate approaches infinity, while their individual relay power goes to zero so that the power emitted per unit

area is constant.

Related Work

Many broadcast algorithms [91] have been ported to wireless networks by abstracting broadcast wireless links as point-to-point, interference free, communication links. While this permitted rapid deployment of wireless networks, ported algorithms often lack desirable attributes like energy or spectral efficiency, or low latency. For instance, simple flooding protocols in wireless networks suffer from the broadcast storm problem [59] resulting from *channel contention* and transmission redundancy. Solutions typically design around the inherent broadcast nature of the wireless link and suppress interference: some obtain collision-free transmissions by heuristically limiting the number of simultaneous transmissions, as in probabilistic flooding [72], others do so by scheduling transmissions [31] or by designating a select group of nodes, like multipoint relaying [69] or dominating connected sets [79], for broadcast.

Instead of adopting a wireless packet-switching approach as above, [90] proposed exploiting the broadcast property of the wireless medium for energy-efficient broadcast by forming a minimum-energy tree rooted at the source node. Similarly, [49] proposed energy-efficient cooperative broadcast protocols but noted that the optimal energy assignment was NP-complete.

Recent work on cooperative transmission, albeit not strictly in the context of broadcasting, has shown that it can significantly improve the system rate, communication range, or power and spectral efficiency [74, 64, 27, 45]. Noting these developments, [33] developed a cross-layer broadcast protocol relying on cooper-

ative diversity which increased network coverage, reduced retransmissions, and decreased latency. However, the protocol requires nodes to know their neighbors and requires explicit coordination among the nodes to achieve cooperation.

[30] and [51, 77] noted that issues of channel contention and transmission redundancy are addressed naturally by cooperative broadcast where transmissions carrying the same information *combine* directly at the physical layer. It was shown that a *single* user can thus flood large networks quickly, subject to power and rate constraints. If the message is transmitted with a power greater than a critical power density then it is guaranteed to reach all network nodes. The other interesting aspect of the analysis in [51, 77] is that it shows how the message is passed from one level of cooperating nodes to the next, by covering a finite region of the network at each passage. The protocols compare favorably to flooding, as shown in [7]. [35, 3, 26] have also investigated techniques that are related to this method.

The analysis in [51, 77], however, does not address the issue of error propagation because the packets are assumed to be long enough that the error can asymptotically be set to zero. Additionally, the question of how best to multiplex several sources, however, is still open. Can we still realize the cooperative gains observed by [77] without scheduling one source at a time? This thesis attempts to address this question. In the same spirit as [77] we do not seek capacity optimal strategies or bounds but we want to rip the benefit of supporting the network flow with relatively well established physical layer strategies. Unlike [77], which considered a single source and so characterized the protocol's success or failure in the deterministic channel through the solution of a recursion on the step-sizes of the regions covered by the broadcast in each time-slot, we show that in the case of multiple sources one must consider a decision process to analyze the cooperative

protocols. The reason is that the propagation of multiple sources in slot $k+1$ depends on the swathes of relays recruited in slot k as well as the interference from other sources.

Literature suggests that decode-and-forward strategies are not always optimal though the argument is not settled for large networks. For example, [5] adopted a quantize-and-forward scheme for achieving capacity to within a constant gap of the cut-set bound for multicast in relay networks. However, this scheme is susceptible to the accumulation of noise, causing the capacity gap to increase with network size. Noise accumulation is also problematic in amplify-and-forward protocols [10] in multi-hop networks. In this light, decode-and-forward schemes are attractive for broadcasting multiple information streams to large networks.

From the point of view of network capacity, coding is also important. The seminal work by Ahlswede et al. [2] proved that network coding provides a capacity advantage in wired networks; [40, 97, 67] extended this idea to analog network coding and physical layer network coding for wireless networks. In other related work, [16] proposed heuristic network coding for MAC layer single-user broadcast and assumed the availability of local topology information. [78] compared network coding to cooperative diversity for single-user broadcast in one-hop networks and showed that network coding yields optimal latency only when broadcasting a long information stream. We incorporate the idea of network coding in our protocols though we do not seek capacity optimal strategies or bounds. We point out, moreover, that our protocols are more general than [40, 97] in that they are applicable in dense networks and account for cooperative relaying and M users.

Another promising approach in MIMO networks is interference alignment (IA) [12, 56], which orthogonalizes the interference at each receiver. It is difficult to

implement in large networks, however, because at each transmitter it requires channel state information (CSI) for all transmitter-receiver pairs and may involve long time delays for extracting full degree of freedom gains. For mitigating long delays, we propose "opportunistic" interference alignment.

The performance of broadcast protocols has traditionally been measured using two metrics - energy consumption [90, 30, 86, 32], or end-to-end latency [24]. Here we evaluate the proposed cooperative protocols in terms of latency.

1.3 Structured Codes for Consensus

Average consensus algorithms [83, 9, 63, 70, 93, 14, 96, 55] provide a robust method for distributed computation in wireless networks lacking central infrastructure. They apply to a wide variety of problems of a distributed nature like information fusion in sensor networks [94], decision making and control among dynamic agents [61], and congestion control in computer and communication networks [83, 19], among others. Average consensus algorithms specify how each node i in a network of N nodes can learn the global average $x_{ave} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i(0)$ by iteratively exchanging messages with their neighboring nodes, where $\mathbf{x}_i(0) \in \mathbb{R}$ are observations known only locally at the start of the distributed computation. The computation speed, or *averaging time* $T_{ave}(\epsilon)$ (c.f. Chapter 4 for definition), of average consensus protocols is typically characterized as the number of iterations required to attain a certain error compared to the actual mean x_{ave}.

There are two types of average consensus algorithms. (1) In *synchronous average consensus*, nodes exchange values with their neighbors on a constant or time-varying graph and there is a common clock that dictates the iterations and

the advance of the algorithm for all nodes; (2) in *asynchronous or randomized average consensus*, progress is driven by random pairwise exchanges between the nodes. Since any node is only allowed to transmit to one neighbor at a time it can be paired naturally with a random access policy. Boyd, et al. [11] have studied the averaging time of both algorithms on general graphs in detail. Several authors [13, 96, 37, 39, 6, 38], among others, have extended the analysis of these algorithms to more sophisticated communication link models with quantization, failures, and/or noise. The impact of fading on these algorithms, however, has not received as much attention.

In asynchronous randomized gossip, nodes activate based on local clocks modeled by a rate μ Poisson process and average pairwise with a randomly chosen neighbor [11] by communicating a packet of duration D. The rate of the clocks is adjusted such that interference from other transmissions that may occur in this window can be neglected. In practice, this constrains the rate of convergence. A simple example can show that link outages triggered by interference increase the averaging time. For the asynchronous Algorithm $\mathcal{A}(P)$ (Lemma 2, [11]) on a graph defined by stochastic matrix $P = P^T$, we can characterize this exactly in the special case that each link is in outage with probability $1 - p$. The network state is updated as $\mathbf{x}(t+1) = W(t)\mathbf{x}(t)$, where W is the mean averaging matrix W. When interference is neglected, $W = (1 - \frac{1}{n})I + \frac{1}{n}P$. Denoting the mean averaging matrix, when link outages are considered, by W', it can be shown that its second largest eigenvalue is $\lambda_2(W') = (1 - \frac{p}{n}) + \frac{p}{n}\lambda_2(P)$. Since $0 \leq p < 1$, we have $\lambda_2(W) \leq \lambda_2(W')$. Therefore, the averaging time must increase since

$$T_{ave}(\epsilon) \leq \frac{3 \log \epsilon^{-1}}{\log \lambda_2(W)^{-1}}. \tag{1.1}$$

We study a more comprehensive example showing that such pairwise averaging consensus algorithms in fading environments that use randomized schedules con-

verge slowly due to interference in Chapter 4. There, we focus on a tractable model of synchronous consensus where time is slotted and a fixed number of nodes pairs average per slot. This model captures the essential features of asynchronous consensus with finite time updates: there, the number of pairwise exchanges per slot would be random. We treat interference in the traditional sense, as a form of degradation of links. We further propose a TDMA schedule that mitigates interference such that if the number of simultaneously active node pairs per slot is the same on average in the TDMA schedule or asynchronous consensus, then the former has more successful pairwise exchanges.

With this motivation, we design *structured channel codes*, employed at the physical layer, for decentralized consensus in large wireless networks. Wireless network gossiping cannot be classified solely as a multiple access or a broadcast problem, it is both. No theoretical result gives the fundamental capacity limits of such a form of communication and there is no previous work indicating the superiority of structured codes in support of this class of network protocols. We review prior relevant work on wireless gossiping for average consensus in Section 5.1. This literature, *e.g.* [11, 21], assumes scheduling with contention and separates source and channel coding. Under a channel contention model, even when optimum scheduling is used, interference and path loss limit network connectivity and affect the speed of convergence to a consensus. Moreover, these works do not account for interference due to fading in wireless networks.

A key notion that this thesis underlines is that in designing networked systems the idea of using *bits* as the currency that abstracts the cost of communications is not the only possibility. A digital communication link is an investment of energy and bandwidth over a certain medium that returns a certain number of bits

that can be spent in conveying the data from point to point in a digitized form. When the sources have common information, the theoretical underpinning for the bandwidth/energy to bit conversion, the separation theorem [17], does not hold. A special case is when the sources are not necessarily correlated but the communication is aimed at calculating a function of the data. In this case, specially structured codes may yield better performance. Notable examples of this are - (1) the construction of Korner and Marton [44] for encoding the modulo-two sum of the binary data and more recently (2) codes by Nazer and Gastpar [57] for computing the sum of analog data.

More examples of structured Medium Access Control (MAC) coding aimed at computing functions of data are offered by the strategies proposed in [53, 54, 47, 4]. Here orthogonal channels are assigned to data types rather than sensors. The multiple access channel models considered are AWGN and i.i.d. fading. The set of orthogonal channels gives the fusion center a scaled histogram of the data types from which one can compute any order invariant function of the data, including the estimate of the mean. If a computation depends on the type-set of the data and not the particular outcome, this strategy scales favorably with the number of sensors. Other work on data gathering illustrates the advantage of a similar principle in the presence of feedback. For example, in [29] nodes are polled about their values and all nodes which have the same answer to the query use the same orthogonal channel dimension to signal it. In this case the channel is again used to compute queries that are function of the data and the sequence of queries is effectively emulating the computation of a codeword compressing the sensor field.

In summary, we propose structured codes for a non-coherent[2] AWGN-MAC

[2]By *non-coherent* we mean an RF channel where neither the transmitter nor the receiver have channel state information, although sufficient time synchronization exists so that the effect of the channel is equivalent to that of a flat fading complex coefficient [68]. The fading is assumed

channel to support the communications for average consensus protocols. These codes result in physical layer cooperative data driven communication.

1.4 Contributions

Most papers that discuss cooperative schemes focus on performance gains in terms of diversity. Our analysis instead reveals that, at a general level, there are additional advantages that stem from employing cooperation at the physical layer rather than using the *collision channel* model of multi-hop networks or the orthogonal relay channels that most other cooperative schemes advocate.

Specifically, in the case of broadcast, our results indicate that in fading channels physical layer decode and forward cooperative broadcast from multiple users is possible without additional overhead, making it a good option for supporting content distribution in wireless networks. Although the analysis focuses predominantly on the two user case, it sheds light on which protocols work and which protocols fail when broadcasting multiple sources. The last go first (LGF) relaying rule shows that in contrast to simply forwarding the sum of received signals as in analog network coding, its better to forward a weighted sum where the signal received weakly is weighted more. Additionally, in the cooperative broadcast of a single source, when transmissions are added up at the physical layer, erroneous transmissions do not necessarily lead to catastrophic error propagation and as the network size grows, the average error can be controlled precisely by controlling the size of the cooperative groups and the power density used in relaying the message. Furthermore, a bandwidth expansion is not necessary to attain this advantage.

to be non-i.i.d.

The required bandwidth can be kept finite by using asynchronous cooperation schemes such as [73] and [52].

In the case of average consensus, we are able to show by using structured codes that - (1) average consensus algorithms in wireless networks are *not intrinsically interference limited*, (2) it is possible to fix the delay per iteration irrespective of the number of nodes, and (3) the features above do not incur extra complexity because the proposed codes can be generated and decoded with extremely simple radio interfaces.

1.5 Outline of the Dissertation

In Chapter 2, we study the simultaneous broadcast of M user messages using physical layer cooperative protocols. We propose and compare four protocols that consider power control and SIC, the idea of network coding, single-user decoding, and interference alignment respectively. Assuming a linear dense network, analysis shows that in deterministic channels, the propagation of messages can be characterized by flows which must cross each other in order for successful broadcast. We find the necessary conditions for this. In fading channels, the flows are probabilistic and the probabilities across time-slots resemble wave fronts.

In Chapter 3, we study the error propagation when cooperatively broadcasting a single source. We derive equations that allow us to evaluate the average BER of uncoded transmission in a dense wireless network with two different cooperative protocols. We prove that *asymptotically, both cooperative protocols have a bounded average BER at a sufficiently high SNR, regardless of the distance covered by the transmission.* Interestingly, the bounds are $\approx Q(\sqrt{SNR})$ for the orthogonal chan-

nel scheme and $\approx 1/SNR$ for the non-orthogonal one. We compare these schemes with each other and with a non-cooperative multi-hop scheme that covers the same distance with the same total power resources. We argue that the cooperative schemes provide better error performance.

In Chapter 4, we argue that standard packet-based average consensus gossiping does not scale in dense wireless networks due to network congestion. We study the effects of interference from fading on the averaging time and total power expenditure by considering the tradeoff between network connectivity and spatial-reuse and show that randomized schedules converge slowly due to interference.

In Chapter 5, we propose a solution to this problem by considering structured codes which utilize cooperative transmission through data driven scheduling. We analyze its performance over an non-coherent AGWN-MAC channel. We characterize the MSE performance and show that adding infinite values with finite precision is possible in spite of the bandwidth limitations.

Finally, Chapter 6 summarizes our results and points to future research directions.

CHAPTER 2

COOPERATIVE BROADCAST OF INFORMATION FROM MULTIPLE SOURCES

In this chapter we consider the problem of multiplexing the information of M independent users who wish to broadcast their messages wirelessly to a dense network. We propose physical layer cooperative decode and forward protocols for this purpose that are immune to the broadcast storm problem. We construct analytical models characterizing the behavior of the cooperative protocols, for deterministic and fading channels, for a network deployed on a line. Special attention is given to the two user case from which we derive necessary conditions to support the information flow of more than two users. It is shown that if nodes are equipped with Successive Interference Cancellation (SIC) receivers and use power control then successful broadcast is achieved in a scalable manner with M, outperforming non-cooperative flooding algorithms in latency. Simulations show that compared to non-cooperative flooding algorithms the cooperative protocols can reduce the latency in flooding $M = 5$ simultaneous users by up to 84%. The protocols also incorporate ideas from analog network coding and interference alignment.

Notation. Matrices are denoted by boldface uppercase letters $\mathbf{U}$, vectors by boldface lowercase letters $\mathbf{h}$. $\mathbf{U}^H$ denotes the hermitian of $\mathbf{U}$. Lowercase letters, typically u, x, denote node coordinates. With some abuse of notation, the node at coordinate x is often referred to simply as node x. Boldface italic upperface letters $\boldsymbol{\mathcal{X}}$ denote sets. $\overline{\boldsymbol{\mathcal{X}}}$ denotes the complement of set $\boldsymbol{\mathcal{X}}$ and $\setminus$ denotes the set difference. $\mathbb{E}$ denotes expectation.

An outline of this chapter is as follows. Section 2.1 lays out the model we study and Section 2.2 introduces four cooperative protocols. Section 2.3 considers de-

terministic channels and determines the conditions necessary for successful broadcast. Section 2.4 extends the analysis to fading channels. Section 2.5 compares the latency of the cooperative protocols to non-cooperative flooding algorithms and provides numerical results supporting our analysis.

2.1 System Model

We consider a linear network, denoted $\boldsymbol{\mathcal{S}}$, where N single-antenna nodes are deployed on a line of length L. The network comprises M source nodes (or users) with mutually independent messages that they wish to broadcast to all nodes in the network. The remaining $N - M$ nodes function as relays. All N nodes in the network are sinks. We assume that all nodes in the network are *reachable* from the M source nodes, not necessarily with a single hop. Let the coordinates of the M source nodes and relays be denoted by $\nu_1, \ldots, \nu_M$, and x respectively. While $\nu_1, \ldots, \nu_M$ are assumed to be fixed, the points x are distributed randomly uniformly in $\boldsymbol{\mathcal{S}}$.

Each source message $W_m = [s_m^1, \ldots, s_m^n]$, $m = 1, \ldots, M$, comprises n symbols s_m^n chosen from some common set $\mathcal{C} = \{c_1, \ldots, c_p\}$. Under the assumption that the bandwidth of the modulation pulse satisfies the Nyquist criterion and that the channel is memoryless, there is no inter-symbol interference. Hence, we can focus on the broadcast of a single symbol s_m per source instead of considering the entire symbol-stream. We refer to the broadcast of a single symbol per source as a broadcast session. The objective is successful broadcast.

Definition 1 (Successful broadcast) *The broadcast session is said to be successful if and only if every node $x \in \boldsymbol{\mathcal{S}}$ has decoded all M user symbols.*

The broadcast session is divided into discrete time-slots. The source nodes initiate the session by simultaneously broadcasting their symbols in the first time slot in a *non-orthogonal* fashion: only a single dimension[1] is used for the transmission of all source symbols. The non-orthogonality of transmissions, in particular, is the essence of *cooperation* in our scheme, as we will see in Section 2.2. The discrete-time baseband signal received by a node at location x from the source nodes in slot 1, post matched filtering, is

$$y_1(x) = \sum_{m=1}^{M} \sqrt{P_s} h(\nu_m, x) s_m + w_1(x), \tag{2.1}$$

where P_s is the common source transmit power, $h(\nu_m, x)$ is the wireless channel coefficient between the source node at ν_m and the relay at x, and $w_1(x)$ is additive white noise which is *i.i.d.* zero mean circularly symmetric complex Gaussian (ZMCSCG) with variance N_o.

We consider two channel models for exposition. The first model considers the wireless links to be *deterministic* coefficients, only accounting for large scale fading or path loss. In this case, the channel between any two nodes located at u and x is given $h(u, x) = \sqrt{\ell(u, x)}$, where $\ell(u, x) = (d_o + |u - x|)^{-2}$ is the path loss with $|u - x|$ the distance between the nodes and d_o a modeling parameter that takes into account the carrier frequency, the scattering environment, and antenna gains.

The second model incorporates fading and assumes the wireless channel to be block faded with frequency-flat slow fading. The channels between any two different node pairs are assumed to be independent. The channel coefficient is ZMCSCG, $h(u, x) \sim \mathcal{CN}(0, \ell(u, x))$ where the variance accounts for the path loss and is *i.i.d.* over each time-slot. We make the following assumption.

[1]This excludes the opportunistic interference alignment protocol which is covered separately in Section 2.2.

A.1. The *receiving* nodes in the network know the channel gains on their *incoming* links through training. This CSI is known without error and is not available at the transmitters.

To decode the received signal (2.1) all N nodes adopt the same strategy. We consider three options. Nodes may use: (1) (**SU**) a single-user receiver, decoding the message embedded in the "strongest" signal; (2) (**SIC**) a successive interference cancellation receiver [82], where the symbols are decoded in the order of decreasing signal strength; or (3) (**SI**) a receiver which cancels interference known through side-information accrued by decoding messages in previous slots [40, 97].

s_m is received at x with a certain signal plus *residual* interference ratio $\mathrm{SINR}_k^m(x)$. The *residual* interference includes terms that depend on the SU, SIC or SI decoding choice. Successful decoding is possible if and only if $\mathrm{SINR}_k^m(x) \geq \tau$, where τ is a pre-defined threshold and $\log_2(1+\tau)$ indicates the rate at which each source symbol is broadcast in bits/s/Hz. We assume-

A.2. SIC and SI receivers cancel interference perfectly.

From an information theoretic standpoint, under certain conditions, successive decoding nearly achieves the Shannon capacity in a multiuser AWGN MAC channel [89, 88]. SIC can also be considered a form of Multi-Packet Reception [80].

Let $\boldsymbol{\mathcal{A}} = 2^{\{1,\ldots,M\}} \setminus \emptyset$ denote the power set of $\{1,\ldots,M\}$ excluding the empty set. We introduce the following definition.

Definition 2 (Decoding Event $\mathcal{E}_k^A$) *The outcome of the decoding operation at each relay x in slot k is denoted by*

$$\mathcal{E}_k^A = \{\text{decoded the set } A \in \boldsymbol{\mathcal{A}}\}.$$

We further assume that-

A.3. Nodes are half-duplex.

A.4. Relay x cooperates in the transmission of s_m during slot k if and only if it decoded s_m for the first time.

Based on A.4 and the definition of $\mathcal{E}_k^A$, regardless of the particular cooperative protocol employed, we make two observations. First, at each time-slot, there is potentially more than one relay forwarding the same source symbol s_m. These relays transmit their symbols simultaneously under the following assumptions.

A.5. All concurrent transmissions throughout the network are synchronized[2] at the symbol level. Furthermore, the signals of all nodes transmitting in the same slot are non-orthogonal, *i.e.*, nodes that transmit simultaneously are *not* allocated separate communication channels (except in OIA).

Second, in each slot $k > 1$, relays forwarding symbol s_m can be grouped into sets (whose memberships vary according to the cooperative protocol) which we also refer to as *levels* as in [77]. We formalize this notion as follows. Based on the decoding outcome $\mathcal{E}_k^A$, we can classify each relay x into one of the following $2^M - 1$ sets:

Definition 3 (Decoding Set $\boldsymbol{\mathcal{S}}_k^A$)

$$\boldsymbol{\mathcal{S}}_k^A = \{x \in \boldsymbol{\mathcal{S}} : \mathcal{E}_k^A \text{ occurred for node } x\},\ A \in \boldsymbol{\mathcal{A}},$$

which permits us to define the set of all relays that *forward* a particular set $A^* \in \boldsymbol{\mathcal{A}}$

[2]While the assumption of network-wide clock synchronization is strong, A.5 is an analytical convenience and not a necessity. In general, quasi-synchronization is a more reasonable assumption. Though one could still use delay diversity methods in that context, the relationship between the received signal and the communication rate becomes cumbersome. The results obtained under the assumption of perfect synchronization can provide benchmarks for the general case.

of user symbols as follows:

Definition 4 (Level Set $\boldsymbol{\mathcal{T}}_k^{A^*}$) *The set of relays that forward the user symbols in set A^* in slot $(k+1)$ is*

$$\boldsymbol{\mathcal{T}}_k^{A^*} = \boldsymbol{\mathcal{S}}_k^{A^*} \setminus (\bigcup_{A:m\in A^*} \boldsymbol{\mathcal{T}}_{k-1}^{A}) \setminus \ldots \setminus (\bigcup_{A:m\in A^*} \boldsymbol{\mathcal{T}}_1^{A}), \tag{2.2}$$

where the exclusion of the sets $\cup_{A:m\in A^*}\boldsymbol{\mathcal{T}}_i^A,\ i=1,\ldots,k-1$, corresponds to the condition that a relay can forward any of the user symbols only once (A.4).

The number of distinct level sets is $|\boldsymbol{\mathcal{A}}| = 2^M - 1$. In general, after k steps of cooperative forwarding, the received signal at a node x is

$$y_{k+1}(x) = \textstyle\sum_{m=1}^{M} \bar{h}_k^m(x) s_m + w_{k+1}(x), \tag{2.3}$$

where $\bar{h}_k^m(x)$ is the *aggregate* channel seen by relay x due to cooperative relaying of the symbol s_m. In the deterministic channel model,

$$\bar{h}_k^m(x) = \sqrt{\textstyle\sum_{A:m\in A} \sum_{u\in \boldsymbol{\mathcal{T}}_k^A} P_r^{u_m} \ell(u,x)}, \tag{2.4}$$

where $P_r^{u_m}$ is the power with which relay u transmits s_m subject to the constraint $\sum_{m=1}^{M} P_r^{u_m} = P_r$, and P_r denotes the total relay power available to each relay for transmission. This model for the channel holds under the assumption:

A.6. In the deterministic channel model *only*, simultaneously transmitting nodes employ a distributed orthogonal space-time code [77] designed for a large number of nodes with the effect that the received power of simultaneously transmitted symbols is equal to the sum of individual powers.

In the fading channel model, on the other hand,

$$\bar{h}_k^m(x) = \textstyle\sum_{A:m\in A} \sum_{u\in \boldsymbol{\mathcal{T}}_k^A} \sqrt{P_r^{u_m}} h(u,x) e^{j\phi_{u_m}}, \tag{2.5}$$

where $\phi_{u_m} \sim U[0, 2\pi]$ is a random phase shift, *i.i.d.* over both the symbol m being forwarded and the relay u.

To characterize the information flow across the network, we focus our efforts on the evolution of the sets $\boldsymbol{\mathcal{T}}_k^A$ over k. Depending on the channel model, the memberships of these sets can be either random or deterministic.

Definition 5 (Information flow) *The progression, over k, of the level sets $\cup_{A:m\in A}\boldsymbol{\mathcal{T}}_k^A$, for a particular user m, is defined as the information flow corresponding to user m.*

In both cases, however, from (2.2), its clear that these sets define a discrete-time dynamical system characterizing the information flow whose solution as $k \to \infty$ indicates whether the broadcast has been successful or not. We examine this in the subsequent sections.

2.2 Cooperative protocols

We focus on simple decode and forward relaying protocols. The protocols do not require relay selection - neither in the sense of global TDMA scheduling [31], nor in the sense of choosing a subset of relays to forward symbols. They also do not require the creation of hierarchies [64]. These features limit overheads.

The decoding outcome $\mathcal{E}_k^A$ is contingent upon the type of receiver employed by the relays. Based on the receiver type, we propose two types of forwarding protocols.

SIC Receiver: In this case, the relays have a variety of forwarding options available

since they may be able to decode a subset of the M sources that were broadcast. These options are captured succinctly by the following protocol.

PC. *(Power Control Protocol)* Power control immediately leads to protocols where relays may choose to forward symbols from all the users they were able to decode or just a subset. Power control is useful since it is known that receiving users with different powers improves the performance of the SIC receiver [88]. Let $\boldsymbol{\mathcal{M}}_k(x) = \cup_{i=1}^{k} \boldsymbol{\mathcal{M}}_i(x)$ denote[3] the history of all the symbols that relay x has decoded till slot k, where

$$\boldsymbol{\mathcal{M}}_k(x) = \{m \in A : \mathcal{E}_i^A \text{occurred for node } x, \forall\ i \leq k, \\ \text{and } m \notin \boldsymbol{\mathcal{M}}_j(x),\ \forall\ j \leq k-1\}.$$

$\boldsymbol{\mathcal{M}}_k(x)$ is also the set of *new* sources that relay x has decoded in slot k. Let P_r denote the total relay power available to each relay for transmission. Then relay x transmits the sum

$$\textstyle\sum_{i \in \boldsymbol{\mathcal{M}}_{k(x)}} \sqrt{P_r^{x_i}} s_i, \text{ subject to} \sum_{i \in \boldsymbol{\mathcal{M}}_{k(x)}} P_r^{x_i} = P_r, \tag{2.6}$$

where $P_r^{x_i}$ is the power with which relay x transmits s_i. Note that $P_r^{x_i}$ is a function of x since relays are allowed to make local decisions regarding the power allocation. The details of the decision process are provided later. In Section 2.4.1, with fading, as given in (2.5), each relay multiplies *each* symbol it forwards by a unit complex number for decorrelating the channels, thus forming a random beam.

Relays using power control basically perform *superposition encoding*. This form of coding is simple in the sense that independent codes can be combined easily at the transmitter to create new codes by using operations like real addition and multiplication. In general though, superposition coding is not optimal for the broadcast channel [82].

[3] $\boldsymbol{\mathcal{M}}_k(x)$ cannot contain repeated elements since elements of a set are unique by definition.

Motivated by recent work on analog network coding [40] and physical-layer network coding [97], we also propose a variation of the above protocol.

PCSI. *(Power Control with Side Information Protocol)* This is similar to protocol PC with the additional assumption that relay x cancels the interference components corresponding to the sources contained in $\boldsymbol{\mathcal{M}}(x)$ from all future received signals. While it improves performance, relays require additional memory resources to store $\boldsymbol{\mathcal{M}}(x)$. Knowledge of the incoming channel gains required for interference cancellation is assured through assumption A.1.

Single-user Receiver: Alternatively, all relays may employ single-user receivers and decode only the source $\hat{m}(x)$ picked according to the following criterion.

SU. *(Single User Receiver Protocol)* In each slot, relays decode the source $\hat{m}(x)$ with the strongest SINR. Relays forward the source $\hat{m}(x)$ they have decoded with power P_r if they have not decoded it before (A.4), *i.e.*, $\hat{m}(x) \in \boldsymbol{\mathcal{M}}_k(x)$. Thus, relay x transmits the signal

$$\begin{cases} \sqrt{P_r}\ s_{\hat{m}(x)}, & \text{if } \hat{m}(x) \in \boldsymbol{\mathcal{M}}_k(x) \\ 0, & \text{otherwise.} \end{cases} \tag{2.7}$$

Motivated by recent work on interference alignment [12], we propose another single-user receiver protocol for use when the channel has fading. It requires $D > 1$ dimensions for each transmission.

OIA. *(Opportunistic Interference Alignment Protocol)* IA requires at least two dimensions (associated to either time-slots or receive antennas, for example). Assuming each transmission now consumes $D < M$ dimensions, in contrast to (2.1),

the length D received vector $\mathbf{y}$ at node x in slot 1 is

$$\mathbf{y}_1(x) = \mathbf{H}_1(x)\mathbf{s} + \mathbf{w}_1(x) \tag{2.8}$$

where $\mathbf{s} = [s_1, \ldots, s_M]^T$, $\mathbf{w}_1(x) \sim \mathcal{CN}(0, N_o I)$, and $\mathbf{H}_1(x)$ is a $D \times M$ matrix obtained by stacking the row vectors $\mathbf{h}^d(x) = \sqrt{P_s}[h^d(\nu_1, x), \ldots, h^d(\nu_M, x)]$, $1 \leq d \leq D$. The channels $h^d(\nu_m, x)$ are *i.i.d.* over d for each fixed m.

OIA opportunistically exploits interference alignment effects that occur during the relaying process - it is likely that some relays will experience a channel matrix $\mathbf{H}_1(x)$ with a high condition number. The key idea is to find the "best" symbol s_m to decode and find beamformers to decode it.

Let $\mathbf{H}_{1,\backslash m}(x)$ denote $\mathbf{H}_1(x)$ without its m^{th} column $\mathbf{h}_1^m \triangleq [h^1(\nu_m, x), \ldots, h^D(\nu_m, x)]^T$. Let $\lambda_D^{\backslash m}$ denote the smallest eigenvalue of $\mathbf{H}_{1,\backslash m}(x)\mathbf{H}_{1,\backslash m}^H(x)$ and $\mathbf{u}_D^{\backslash m}$ the associated eigenvector. Then, x decodes the symbol $\hat{m}$ where

$$\hat{m} = \arg\max_m \frac{|(\mathbf{u}_D^{\backslash m})^H \mathbf{h}_1^m|^2}{N_o + \lambda_D^{\backslash m}}. \tag{2.9}$$

The relay x then transmits the same signal as in (2.7). In subsequent time-slots, $\mathbf{H}_k(x)$ comprises realizations of aggregate channels $\bar{h}_k^m(x)$ instead (c.f. Section 2.4).

OIA requires more sophisticated receivers than protocol SU but it yields better performance because $\mathbf{u}_D^{\backslash m}$ projects $\mathbf{h}_1^m$ along the direction in which the interference is the least. Further, compared to SU, $\lambda_D^{\backslash m}$ is no bigger than the smallest diagonal element of $\mathbf{H}_{1,\backslash m}(x)\mathbf{H}_{1,\backslash m}^H(x)$, *i.e.*, $\lambda_D^{\backslash m} \leq \min_d \sum_{n \neq m} |h^d(\nu_n, x)|^2$. The improvement OIA provides over SU is examined numerically in Section 2.5. Moreover, since interference alignment effects are opportunistic, we do not require waiting for long durations to see gains, unlike [56].

2.3 Deterministic Channel

Studying the protocols in deterministic channels is helpful in understanding their fundamental behavior. The general problem with M source nodes, however, has unwieldy combinatorial aspects, especially for protocols PC and PCSI. For tractability, we restrict our analysis to $M = 2$ source nodes, located at coordinates $\nu_1 = 0$ and $\nu_2 = L$ respectively. This simplification allows us to establish a base case. Increased interference from multiple users implies that the performance of the broadcast protocols in the general case can be no better. There is a connection as well to two-way relay strategies [18], that we here generalize in a multi-relay multi-level setting.

For $M = 2$, $\boldsymbol{\mathcal{A}} = \{\{1\}, \{2\}, \{1, 2\}\}$. So, with abuse of notation, we are interested in characterizing three level sets: $\boldsymbol{\mathcal{T}}_k^1, \boldsymbol{\mathcal{T}}_k^2$, and $\boldsymbol{\mathcal{T}}_k^b$, where $b = \{1, 2\}$. They denote the sets of relays that forward symbols s_1 or s_2 exclusively, or both symbols, s_1 and s_2, respectively. The two flows start at opposite edges of the network and move towards each other. Intuitively, for successful broadcast, these flows must *cross* each other at some coordinate ζ.

Definition 6 (Flow crossing) *Two information flows m and m', $m \neq m'$, cross each other in slot k^* if* $\max \cup_{A:m\in A} \boldsymbol{\mathcal{T}}_{k^*}^A \leq \zeta \leq \min \cup_{A':m'\in A'} \boldsymbol{\mathcal{T}}_{k^*}^{A'}$ *and* $\min \cup_{A:m\in A} \boldsymbol{\mathcal{T}}_{k^*+1}^A \geq \zeta \geq \max \cup_{A':m'\in A'} \boldsymbol{\mathcal{T}}_{k^*+1}^{A'}$.

Since the elements of $\boldsymbol{\mathcal{T}}_k^A$ can be ordered, $\max \boldsymbol{\mathcal{T}}_k^A$ and $\min \boldsymbol{\mathcal{T}}_k^A$ are well-defined. Crossing of flows is a necessary condition for successful broadcast but insufficient.

2.3.1 Power Control Protocol (PC)

If a relay x only decodes a single user then it forwards that user's symbol with power P_r. On the other hand, if x is able to decode both users, then x allocates power such that $P_r^{x_m} = \eta P_r$ and $P_r^{x_n} = (1-\eta)P_r$, $m, n \in \{1,2\}$, $m \neq n$ (Table 2.1). Coefficient $\eta \in [0,1]$ is chosen locally (see (2.6)) and modulates the power allocated for relaying each user's symbol.

EQUAL POWER ALLOCATION

We examine the protocol under the simplifying assumption of *equal* power allocation ($\eta = \frac{1}{2}$) first for understanding. To determine the level sets $\boldsymbol{\mathcal{T}}_k^i$, $i = 1, 2, b$, one needs to know which relays have decoded which symbols: relay x first checks whether it can decode the symbol s_m it receives with the greatest power: $\mathrm{SINR}_k^m(x) \geq \tau$. If it is able to decode s_m, it cancels the interference s_m generates and checks whether it can decode the other symbol: $\mathrm{SNR}_k^n(x) \geq \tau, m, n \in \{1,2\}$, $n \neq m$. The net effect is that relay x can either decode only symbol s_m, or both symbols s_m and s_n. Thus, we need to solve the system:

$$\mathrm{SINR}_k^m(x) = \frac{\sum_{u \in \boldsymbol{\mathcal{T}}_{k-1}^m \cup \boldsymbol{\mathcal{T}}_{k-1}^b} P_r^u \ell(u,x)}{N_o + \sum_{u \in \boldsymbol{\mathcal{T}}_{k-1}^n \cup \boldsymbol{\mathcal{T}}_{k-1}^b} P_r^u \ell(u,x)} \gtrless \tau, \tag{2.10}$$

$$\mathrm{SNR}_k^n(x) = N_o^{-1} \sum_{u \in \boldsymbol{\mathcal{T}}_{k-1}^n \cup \boldsymbol{\mathcal{T}}_{k-1}^b} P_r^u \ell(u,x) \lessgtr \tau, \tag{2.11}$$

Table 2.1: Power Control in Protocol PC

s_1	s_2	Condition
$\sqrt{\eta P_r}$	$\sqrt{(1-\eta)P_r}$	if x decoded s_2 first
$\sqrt{(1-\eta)P_r}$	$\sqrt{\eta P_r}$	if x decoded s_1 first

$m, n \in \{1, 2\}$, $m \neq n$, where the signal power relay x receives from user m is computed using (2.4). Solving (2.10) and (2.11) for x is complicated.

Remark 1 *The information flow dynamics for this case and even the general case with M users for a finite sized network can be obtained by solving the system* (2.10)-(2.11) *numerically. We also show the broadcast efficacy of the protocols through simulation in Section 2.5 for networks with up to $M = 5$ users.*

To characterize the sets $\boldsymbol{\mathcal{T}}_k^i$ analytically we resort to asymptotics. We let the network become infinitely dense by letting the number of nodes $N \to \infty$. This technique has also been used in [77, 81] to wash out the randomness associated with the random node deployment in the network. Concurrently, we let the relay transmit power vanish, $P_r \to 0$, so that the relay transmit power per unit area remains constant $P_r \rho \to \text{const}$, where $\rho = \frac{N}{L}$ is the node density of the network. Using Theorem 1, from [77], we know that in the limit $N \to \infty, P_r \to 0$, for $m = 1, 2$,

$$\sum_{u \in \boldsymbol{\mathcal{T}}_{k-1}^m \cup \boldsymbol{\mathcal{T}}_{k-1}^b} P_r^u \ell(u, x) \to \int_{\boldsymbol{\mathcal{T}}_{k-1}^m \cup \boldsymbol{\mathcal{T}}_{k-1}^b} P_r^u \rho \ell(u, x) du. \tag{2.12}$$

Intuitively, every infinitesimal area du has ρdu nodes that transmit. This transforms our discrete-valued combinatorial problem to a real-valued, though non-linear, one. In this infinitely dense regime, the level sets $\boldsymbol{\mathcal{T}}_{k-1}^m$ are, in general, unions of continuous intervals. Even with this simplification, (2.10)-(2.11) is a system comprising high order equations in x that remains difficult to solve. It turns out, however, we can show there is a structure to the level sets and their evolution.

If we solve the system (2.10)-(2.11) numerically, as shown in Fig. 2.1 for one choice of network parameters, level sets for both users initially form at their respective source locations, 0 and L. Then the sets progressively move toward each

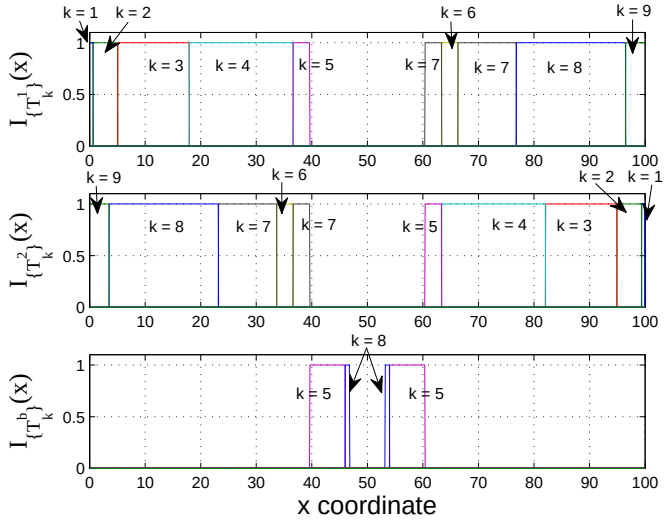


Figure 2.1: The initial level sets $\boldsymbol{\mathcal{T}}_1^1$ and $\boldsymbol{\mathcal{T}}_1^2$ form at 0 and $L = 100$ respectively and then begin moving toward each other. The level membership of each node x is indicated by $I_{\{\boldsymbol{\mathcal{T}}_k^i\}}(x)$, $i = 1, 2, b$, where $I_{\{\cdot\}}(\cdot)$ denotes the indicator function. Level memberships across time slots are color coded. The figure was generated by solving the system (2.10)-(2.11) numerically.

other before crossing over at $\zeta = \frac{L}{2}$ and continuing to L and 0 respectively. In general, in each slot k, the level sets belong to one of four *network states*, denoted by $\boldsymbol{\mathcal{L}}_k$. Furthermore, given the current state $\boldsymbol{\mathcal{L}}_k$ we can predict the next state $\boldsymbol{\mathcal{L}}_{k+1}$, prior to the crossing of flows. The proof of these claims constitutes the main result of this section.

First consider the network state immediately after the original broadcast of symbols in $k = 0$. Since an SIC receiver decodes signals in order of decreasing strength and both source transmit powers are identical, all relays $x \leq \frac{L}{2}$ decode s_1 first while relays $x > \frac{L}{2}$ decode s_2 first. By definition, $\boldsymbol{\mathcal{T}}_1^1$ is the set of all nodes that decode *only* symbol s_1, *i.e.*, $x \in \boldsymbol{\mathcal{T}}_1^1$ only if $\text{SINR}_1^1(x) \geq \tau \ \cap \ \text{SNR}_1^2(x) < \tau$. So, $x \in \boldsymbol{\mathcal{T}}_1^1$ only if $x \leq \frac{L}{2}$. The precise characterization of the level sets in $k = 1$ is given by the following Lemma.

Lemma 1 *When x_0^* and $\bar{x}_0$ denote the respective solutions of* $\mathrm{SINR}_1^1(x) = \tau$ *and* $\mathrm{SNR}_1^2(x) = \tau$*, which we find to be*

$$x_0^* \geq \left(\frac{\tau N_o}{P_s} + (\frac{1+\sqrt{\tau}}{L+2d_0})^2 \right)^{-\frac{1}{2}} - d_0, \;\; \bar{x}_0 = L + d_0 - \sqrt{P_s/\tau N_o},$$

the network state $\boldsymbol{\mathcal{L}}_1$ satisfies the following statements:

(i) If no feasible x_0^* exists in the interval $(0, \frac{L}{2}]$ then the flows stop and no further transmissions occur.

(ii) Otherwise, the network state $\boldsymbol{\mathcal{L}}_1$ is $\boldsymbol{\mathcal{L}}_1^{(1)}$ if $\bar{x}_0 > x_0^*$, $\boldsymbol{\mathcal{L}}_1^{(2)}$ if $\bar{x}_0 < 0$, and $\boldsymbol{\mathcal{L}}_1^{(3)}$ if $\bar{x}_0 < x_0^*$, where the three network states $\boldsymbol{\mathcal{L}}_1^{(1)}$, $\boldsymbol{\mathcal{L}}_1^{(2)}$, and $\boldsymbol{\mathcal{L}}_1^{(3)}$ are defined in Fig. 2.3.

Proof First note that the level sets must be symmetric about $\frac{L}{2}$ (c.f. Fig. 2.3). To see why, note that $\mathrm{SINR}_1^1(x) = \mathrm{SINR}_1^2(L-x)$ because of symmetry in the path loss function $\ell(u,x) = \ell(L-u,x)$ and equal transmit power allocation. Similarly $\mathrm{SNR}_1^1(x) = \mathrm{SNR}_1^2(L-x)$. Therefore, it is sufficient to consider $x \in (0, \frac{L}{2}]$.

The proof is simple. We show that $\mathrm{SINR}_1^1(x) = \tau$ and $\mathrm{SNR}_1^2(x) = \tau$ have unique solutions x_0^* and $\bar{x}_0$ since path-loss $\ell(u,x)$ is monotonic. The solutions of $\mathrm{SINR}_1^1(x) \geq \tau$ and $\mathrm{SNR}_1^2(x) < \tau$ are therefore contiguous intervals. Comparing these intervals results in the Lemma.

Let $\theta(x) = (x + d_o)^{-2}$ and $\psi(x) = \tau(d_0 + L - x)^{-2}$; these functions are shown in Fig. 2.2 and satisfy-

(i) $\theta''(x) > 0$ and $\psi''(x) > 0$, $\forall x$, so they are convex.

(ii) $\lim_{x\to\infty} \theta(x) = 0$ and $\lim_{x\to-\infty} \psi(x) = 0$.

(iii) $\theta'(x) < 0$ and $\psi'(x) > 0$, $\forall x \in (0, L)$, so $\theta(x)$ and $\psi(x)$ are monotonically decreasing and increasing respectively in this interval.

We find a unique feasible solution x_0^* in the range of interest $(0, \frac{L}{2}]$. Recall,

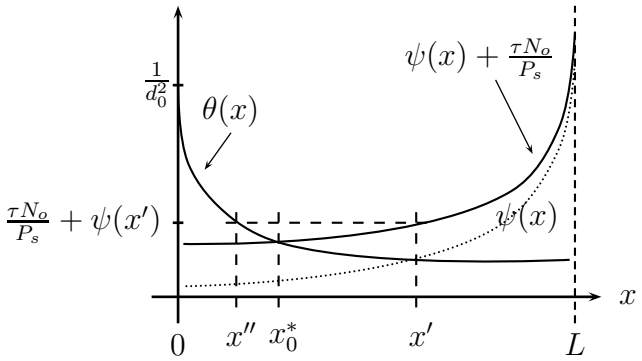


Figure 2.2: x'' is the lower bound to the exact solution, x_0^*, of the quartic equation in (2.13). The figure is not to scale.

$\text{SINR}_1^1(x) = \frac{P_s \ell(0,x)}{N_o + P_s \ell(L,x)}$. Using the definitions of the path loss and the functions defined above, we can rewrite $\text{SINR}_1^1(x) = \tau$ as

$$\theta(x) \;\; = \;\; \tau N_o/P_s + \psi(x), \tag{2.13}$$

which is a quartic equation in x. We bound its solution.

From properties (i) - (iii), $\theta(x) - \psi(x) = 0$ can have at most one real positive root. The solution of $\theta(x) - \psi(x) = 0$, is $x' = [L + d_o(1 - \sqrt{\tau})](1 + \sqrt{\tau})^{-1}$. Now consider (2.13). The function $\tau N_o/P_s + \psi(x)$ satisfies properties (i) and (iii) and $\lim_{x \to -\infty} \tau N_o/P_s + \psi(x) = \tau N_o/P_s$. Therefore, (2.13) must have a unique real positive root x_0^* in $(0, \frac{L}{2}]$ such that $x_0^* \leq x'$. Using property (iii), x_0^* can be lower bounded as

$$x_0^* \geq x'' = \theta^{-1}\left(\tau N_o/P_s + \psi(x')\right), \tag{2.14}$$

which can be simplified to $x_0^* \geq (\frac{\tau N_o}{P_s} + (\frac{1+\sqrt{\tau}}{L+2d_0})^2)^{-\frac{1}{2}} - d_0$, which implies $\text{SINR}_1^1(x) \geq \tau$ if $x \leq x_0^*$.

By definition, $\boldsymbol{\mathcal{T}}_1^b$ is the set of all nodes that decode *both* s_1 and s_2. Equivalently, $x \in \boldsymbol{\mathcal{T}}_1^b$ if $(\text{SINR}_1^1(x) \geq \tau \;\cap\; \text{SNR}_1^2(x) \geq \tau) \bigcup (\text{SINR}_1^2(x) \geq \tau \;\cap\; \text{SNR}_1^1(x) \geq \tau)$, where the union of events accounts for the possibility that the SIC receiver at x

may decode either s_1 or s_2 first depending on which source it is physically closer to.

We have already solved for x: $\text{SINR}_1^1(x) \geq \tau$. Noting that $\text{SNR}_1^1(x) = P_s\ell(L,x)/N_o$, one can verify that for $x \leq \frac{L}{2}$, $\text{SNR}_1^2(x) < \tau$ only if

$$x \leq \bar{x}_0 = L + d_0 - \sqrt{P_s/\tau N_o}. \tag{2.15}$$

We can determine the network state $\boldsymbol{\mathcal{L}}_1$ by comparing x_0^* and $\bar{x}_0$. If $\bar{x}_0 > x_0^*$, then no nodes are able to decode both sources and $\boldsymbol{\mathcal{T}}_1^b = \emptyset$. In this case, $r_1 = x_0^*$, leading to $\boldsymbol{\mathcal{L}}_1^{(1)}$. If $r_0 < \bar{x}_0 < x_0^*$, then some nodes are able to decode both sources and $\boldsymbol{\mathcal{T}}_1^b = (r_1, q_1] \cup [L - q_1, L - r_1)$ where $r_1 = \bar{x}_0$ and $q_1 = x_0^*$, leading to $\boldsymbol{\mathcal{L}}_1^{(3)}$. *The union with the interval* $[L - q_1, L - r_1)$ *follows by repeating the above process for relays* $x > \frac{L}{2}$. Finally, if $\bar{x}_0 \leq r_0 = 0$, then $\boldsymbol{\mathcal{T}}_1^1 = \boldsymbol{\mathcal{T}}_1^2 = \emptyset$ and $\boldsymbol{\mathcal{T}}_1^b = (0, r_1] \cup [L - r_1, L)$, leading to $\boldsymbol{\mathcal{L}}_1^{(2)}$. ■

At this point, we note that given the current the network state $\boldsymbol{\mathcal{L}}_k$, the state in the next time-slot $k+1$ is known through Lemmas 2 - 4. We note the following before stating the Lemmas.

Remark 2 *Network states in consecutive slots,* e.g.*,* $\boldsymbol{\mathcal{L}}_1$ *and* $\boldsymbol{\mathcal{L}}_2$*, touch each other, as shown in Fig. 2.1. This is a consequence of the monotonicity of the path-loss function* $\ell(u,x)$ *which guarantees that the solutions of* $\text{SINR}_2^1(x) = \tau$ *and* $\text{SNR}_2^2(x) = \tau$ *are unique. An implication of this is that as the two information flows move toward the center, they cover all nodes in their paths.*

Let r_k denote the right boundary of the level set $\boldsymbol{\mathcal{T}}_k^1$ with initialization $r_0 = 0$.

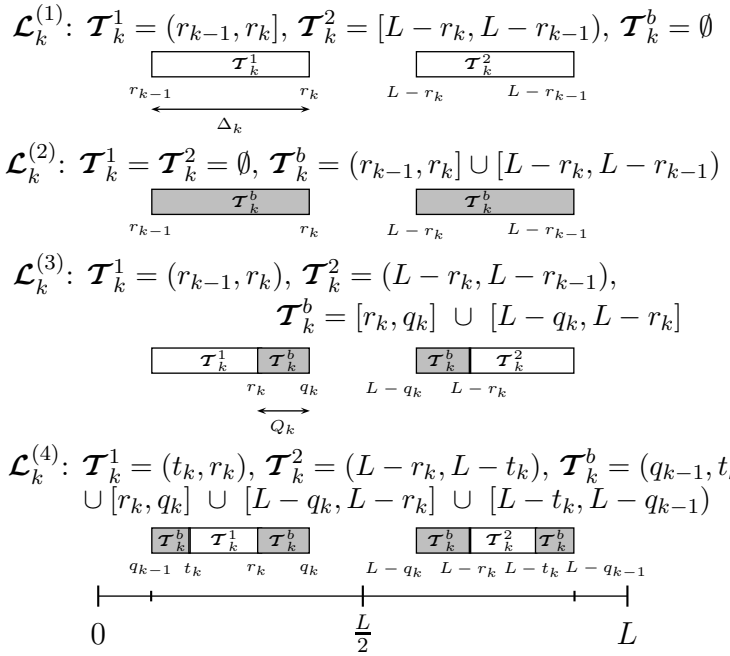


Figure 2.3: Four network states are possible, as specified in Theorem 1. For clarity, intervals on the real line are represented as rectangles.

Lemma 2 *Let $\Delta_k = r_k - r_{k-1}$. If the current network state is $\boldsymbol{\mathcal{L}}_k^{(1)}$, the next state is determined by the following criteria.*

(a) The information flows must be contained within the open interval $(r_k, L - r_k)$.

(b) The solutions of $\mathrm{SINR}_{k+1}^1(x) = \tau$ and $\mathrm{SNR}_{k+1}^2(x) = \tau$ are x_k^* and $\bar{x}_k$ where

$$\begin{aligned} x_k^* &\geq \frac{r_k + r_{k-1}}{2} + \frac{1}{2}\sqrt{\Delta_k^2 + 4\left(\frac{\tau N_0}{P_r \rho \Delta_k} + \psi(x_k')\right)^{-1}} - d_0, \\ \bar{x}_k &= L - \frac{r_k + r_{k-1}}{2} - \frac{1}{2}\sqrt{\Delta_k^2 + 4P_r\rho\Delta_k/\tau N_0} + d_0, \end{aligned}$$

where $\psi(x) \triangleq \tau(L - x + d_0 - r_k)^{-2}$, $\theta(x) \triangleq (x + d_o - r_k)^{-1}(x + d_o - r_{k-1})^{-1}$ and x_k' is the solution of $\theta(x) = \psi(x)$. Then,

$$\boldsymbol{\mathcal{L}}_{k+1} = \begin{cases} \boldsymbol{\mathcal{L}}_{k+1}^{(1)}, & \text{if } \bar{x}_k > x_k^*, \\ \boldsymbol{\mathcal{L}}_{k+1}^{(2)}, & \text{if } \bar{x}_k < r_k, \\ \boldsymbol{\mathcal{L}}_{k+1}^{(3)}, & \text{if } r_k < \bar{x}_k < x_k^*. \end{cases} \tag{2.16}$$

(*c*) *If no feasible* x_k^* *exists in the interval* $(r_k, \frac{L}{2}]$ *then the flows stop and no further transmissions occur.*

Proof See Appendix A.1. ■

Lemma 2 generalizes the result of Lemma 1 to an arbitrary slot $k \leq k^*$. Statement (*b*) implies that if the current network state is $\boldsymbol{\mathcal{L}}_k^{(1)}$ then the information flows must always move toward the center of the network or stop completely. They cannot cross each other from state $\boldsymbol{\mathcal{L}}_k^{(1)}$.

Remark 3 $\exists$ *critical threshold* τ^* s.t. *if* $\tau > \tau^*$ *then* $\boldsymbol{\mathcal{L}}_k = \boldsymbol{\mathcal{L}}_k^{(1)}$ *with* $\Delta_{k+1} < \Delta_k, \ \forall k > 0$, i.e., *the flows terminate without ever crossing.*

Lemma 3 *If* $\tau > 1$ *and the current network state is* $\boldsymbol{\mathcal{L}}_k^{(2)}$ *then all transmissions cease in slot* $k+1$.

Proof Consider any $x \in \boldsymbol{\mathcal{S}}$. From the definition of $\boldsymbol{\mathcal{L}}_k^{(2)}$ in Fig. 2.3, x receives either symbol with power $P_r\rho/2 \int_{r_{k-1}}^{r_k} \ell(u,x)du + P_r\rho/2 \int_{L-r_k}^{L-r_{k-1}} \ell(u,x)du$. From (2.10), this means $\text{SINR}_{k+1}^m(x) < \tau$ for any $m \in \{1,2\}$, provided $\tau > 1$. Since this is true $\forall\ x \in \boldsymbol{\mathcal{S}}$, all flows terminate. ■

We have shown so far that states $\boldsymbol{\mathcal{L}}_k^{(1)}$ and $\boldsymbol{\mathcal{L}}_k^{(2)}$ do not permit the information flows to cross. Unless the flows cross, which can happen only from states $\boldsymbol{\mathcal{L}}_k^{(3)}$ or $\boldsymbol{\mathcal{L}}_k^{(4)}$, as we now show, the broadcast cannot succeed.

Lemma 4 *If the current network state is* $\boldsymbol{\mathcal{L}}_k^{(3)}$, *the following statements hold:*
(*a*) *Using an asymptotic approximation, under the condition*

$$\frac{1-\tau}{2d_0} < \frac{\tau N_o}{P_r\rho} + \frac{r_k\tau}{(L-q_k+2d_0)(L-q_k-r_k+2d_0)} - \frac{r_k}{q_kQ_k},$$

the next state $\boldsymbol{\mathcal{L}}_{k+1}$ is one of types defined in Fig. 2.3 and is contained in the interval $(q_k, L - q_k)$.

(b) Let $Q_k = q_k - r_k$ (see Fig. 2.3). The information flows cross if the following condition is satisfied:

$$P_r \rho(\Delta_k + Q_k) > \tau N_0 (L - 2r_{k-1} + d_0)^2. \tag{2.17}$$

A similar result holds for state $\boldsymbol{\mathcal{L}}_k^{(4)}$, with modified conditions.

Proof See Appendix A.2. ■

When both statements of Lemma 4 hold at k^* then we say that the level sets have *split*, *i.e.*, they are no longer contiguous. When this happens, it becomes difficult to characterize $\boldsymbol{\mathcal{L}}_{k+1}$ because finding the boundaries becomes non-trivial.

We summarize the results presented so far in Theorem 1 below. It gives the necessary conditions for the crossing of information flows in slot k^* and shows how the flows evolve up till that point.

Theorem 1 *Consider protocol PC with equal power allocation and assume $\tau > 1$. The following statements hold $\forall\ k \leq k^*$, where k^* is the slot in which the information flows cross:*

(i) Network state $\boldsymbol{\mathcal{L}}_k$ must be one of the four types in Fig. 2.3.

(ii) As the two information flows move toward the center they cover all nodes in their paths.

(iii) Only states $\boldsymbol{\mathcal{L}}_k^{(3)}$ and $\boldsymbol{\mathcal{L}}_k^{(4)}$ permit the flows to cross.

Proof *(i)* The first network state is $\boldsymbol{\mathcal{L}}_1^{(1)}$-$\boldsymbol{\mathcal{L}}_1^{(3)}$ by Lemma 1. By Lemma 2, $\boldsymbol{\mathcal{L}}_k^{(1)}$ can only be followed by $\boldsymbol{\mathcal{L}}_{k+1}^{(1)}$-$\boldsymbol{\mathcal{L}}_{k+1}^{(3)}$. Lemma 3 showed that flows die after $\boldsymbol{\mathcal{L}}_k^{(2)}$.

Finally, Lemma 4 showed that $\boldsymbol{\mathcal{L}}_k^{(3)}$ can lead to any of the four possible states. The same holds for $\boldsymbol{\mathcal{L}}_k^{(4)}$.

(*ii*) This is a consequence of the monotonicity of the path-loss function $\ell(u, x)$ which guarantees that the solutions of $\text{SINR}_k^m(x) = \tau$ and $\text{SNR}_k^m(x) = \tau$, $m \neq n \in \{1, 2\}$ are unique.

(*iii*) Follows from Lemmas 1 - 4. ■

It is easy to see that because of equal power allocation, the level sets $\boldsymbol{\mathcal{T}}_k^1$ and $\boldsymbol{\mathcal{T}}_k^2$ must remain symmetric $\forall k$, even after the information flows cross.

VARIABLE POWER ALLOCATION

We now discuss protocol PC with $\eta \in [0, 1]$. Note that until the joint level sets $\boldsymbol{\mathcal{T}}_k^b$ form no power control is invoked and the evolution of level sets remains identical to Section 2.3.1. It is known that SIC receivers perform better when the power levels of the received signals are easily distinguishable. Power allocation helps with this. Although the optimum η is hard to compute analytically[4], we show that unlike Lemma 3, under certain conditions flows can cross from state $\boldsymbol{\mathcal{L}}_k^{(2)}$ for some $\eta > 0.5$.

Lemma 5 *Let $\Delta_k = r_k - r_{k-1}$. If the current network state is $\boldsymbol{\mathcal{L}}_k^{(2)}$, the information flows can cross if $\eta = 1$ and*

$$\tau(L - 2r_{k-1} + d_0)^{-2} \leq (d_0 + \Delta_k)^{-2} - \tau N_o (P_r \rho \Delta_k)^{-1}. \tag{2.18}$$

[4]The computational complexity of the power allocation problem is independent of the network size N and number of users M since η is chosen locally (Table 2.1).

Proof See Appendix A.2. ■

The condition above ensures that the interference from the transmission of symbol s_m is low enough to permit symbol s_n, $m \neq n$, to be decoded. While Lemma 5 does not guarantee that flows will cross, the possibility of it happening is interesting.

Intuitively, relays should allocate more power for the transmission of the symbol received feebly (*i.e.*, decoded second) because its information flow requires assistance in crossing. We call this the *last go first* (LGF) transmission rule. From Table 2.1, this suggests an optimal value of $\eta > 0.5$ and this is borne out in simulations (c.f. Fig. 2.4(a)). In Section 2.5 we numerically show the optimal value of η approaches 1. LGF shows that in contrast to simply forwarding the sum of received signals as in analog network coding, its better to forward a weighted sum where the signal received weakly is weighted more.

2.3.2 Power Control with Side-Information Protocol (PCSI)

Relays that have already decoded the symbol of at least one user can exploit this side-information and cancel out interference components corresponding to these decoded users from all future received signals. In the case of two users, as soon as a relay has decoded the symbol of one user, it no longer experiences any interference in decoding the second user. The subsequent set memberships of such a relay are analogous to the evolution of level sets when a single user broadcasts its symbol, as addressed in [77]. Intuitively, this makes it much easier for the information flows to cross, improving the performance of the broadcast protocol.

No relays can have the opportunity of using side information for $k < 2$ so Lemma 1 holds. For $k \geq 2$, the following result shows that in contrast to Theorem 1, it is possible for the flows to cross from *each* network state.

Lemma 6 *The use of side-information by the nodes enables the information flows to cross from each of the network states defined in Theorem 1.*

Proof See Appendix A.3. ■

Since the use of side-information allows the information flows to cross from each network state, it can only improve the broadcast latency compared to protocol PC.

2.3.3 Extension to M Users

Due to additional interference, it is not straightforward to extend the analytical characterization of the level sets to an arbitrary number of users in a linear network. However, one can formulate a necessary condition for the successful broadcast of M users based on the conditions derived earlier for the crossing of information flows of two users.

Lemma 7 *The messages of M users in a linear network can be broadcast successfully only if the information flows of any two users, considered as a pair while neglecting the interference from all other users, can cross each other. This provides a necessary condition. There are $\binom{M}{2} = \frac{M(M-1)}{2}$ possible pairs of users that one must check.*

This necessary condition is applicable to both protocols PC and PCSI. Note that the stated condition is not sufficient for successful broadcast.

2.3.4 Single-User Receiver Protocol (SU)

In this protocol, relays use single-user receivers and simply decode and forward the source symbol that they receive with the greatest power, provided that they have not already forwarded that source before. Relays do not exploit side-information. Surprisingly, under the deterministic channel model, this protocol does not support successful broadcast of even two users' messages.

Lemma 8 *Under the deterministic channel model, cooperative protocol SU does not permit the crossing of information flows. Therefore, it cannot support successful broadcast.*

Proof The proof is straightforward. If single-user receivers are used, $\boldsymbol{\mathcal{T}}_k^b = \emptyset, \forall k$, *i.e.*, relays cannot decode more than one source symbol in a given slot. Therefore, the network state must be $\boldsymbol{\mathcal{L}}_k^{(1)}, \forall k$. However, in Lemma 2 we showed that the information flows can never cross from state $\boldsymbol{\mathcal{L}}_k^{(1)}$. Since the flows cannot cross, the broadcast session must terminate unsuccessfully. ■

We will see, in Section 2.4.3, where we analytically characterize protocol SU for the general case of M users, that it does not guarantee successful broadcast even under a channel model that incorporates fading.

2.4 Fading Channel

We now consider a channel model with fading (c.f. Section 2.1). Before analyzing the cooperative protocols in depth, we discuss how channel randomization induced by fading allows some "lucky" relays to decode source messages that they would

not have been able to decode if the channel was deterministic. Such opportunistic gains have been noted in prior literature, *e.g.* [87].

The opportunistic gains are demonstrated most easily by considering both M and N in the limit while keeping the network size fixed. Assume the source nodes are located randomly in $\boldsymbol{\mathcal{S}}$. In the deterministic model without fading, it is clear that in the first time slot $\lim_{M\to\infty} \mathrm{SINR}_1^m(x) = 0$, $\forall\ m$, and $\forall x \in \boldsymbol{\mathcal{S}}$, because the interference grows unboundedly. This means none of the messages broadcast by the source nodes will be decoded.

In contrast, in the fading model, assuming all channels $h(\nu_m, x)$ are *i.i.d.* standard complex normal and neglecting path-loss for a moment, each source node is able to recruit multiple relays for forwarding its message, as shown by the following lemma.

Lemma 9 *Under protocol PC, when all channels are i.i.d., each source node is able to recruit* $\Theta(\log N)$ *relays to help forward its message in the first time slot* $k = 1$ *in the limit of large* N *provided* M *scales as* $O(\log N)$.

Proof See Appendix A.4. ■

It is worth noting that these opportunistic gains will arise in subsequent time slots and for finite M as well although the gains may be diminished (reflected in the number of relays recruited per source) when we include path-loss considerations. Thus, we expect the cooperative protocols to exhibit lower broadcast latency when the channel is fading impaired.

In order to characterize the information flow across the network, we study the evolution of the level sets $\cup_{A:m\in A} \boldsymbol{\mathcal{T}}_k^A$ over k. In any given instance of the protocol,

however, the memberships of these sets are *random*. Accordingly, define-

Definition 7 (Level probability $P_k^{A^*}(x)$) *The level probability $P_k^{A^*}(x)$ is defined on a* per node *basis. It is the probability that relay x belongs to level set $\boldsymbol{\mathcal{T}}_k^{A^*}$ in slot k: $P_k^{A^*}(x) = P(x \in \boldsymbol{\mathcal{T}}_k^{A^*})$.*

The system dynamics, which are now stochastic, are captured by these level probabilities. The following simplification will be used in the sequel.

$$\begin{aligned} P_k^{A^*}(x) = P(x \in \boldsymbol{\mathcal{T}}_k^{A^*}) &\stackrel{(a)}{=} P\left[x \in \boldsymbol{\mathcal{S}}_k^{A^*} \cap_{i=1}^{k-1} \left[\cap_{A:m\in A^*} \bar{\boldsymbol{\mathcal{S}}}_i^{A}\right]\right] \\ &\stackrel{(b)}{=} P(x \in \boldsymbol{\mathcal{S}}_k^{A^*}) \prod_{i=1}^{k-1} P(x \in \cap_{A:m\in A^*} \bar{\boldsymbol{\mathcal{S}}}_i^{A}), \end{aligned} \tag{2.19}$$

where (a) follows by substituting the definition of $\boldsymbol{\mathcal{T}}_i^{A}$ and applying the distributive law of set operations recursively to (2.2) and (b) follows because, as we will see in each of the protocols, the aggregate channel gains $|\bar{h}_i^m(x)|^2$ are independent over i (A.4). The latter product of k terms in (2.19) accounts for assumption A.4: if x has forwarded any symbol s_m for $m \in A^*$ in prior slots then $x \notin \boldsymbol{\mathcal{T}}_k^{A^*}$.

Computing $P_k^{A^*}(x)$ for arbitrary M is not straightforward for protocols PC and PCSI - they utilize SIC receivers so there will be $2^M - 1$ different level probabilities to compute. We will restrict attention to $M = 2$ for protocols PC and PCSI. Protocol SU, in contrast, only requires M level probabilities and we are able to analyze it for M users.

2.4.1 Power Control Protocol (PC)

Recall the aggregate channel from (2.5). It is Gaussian:

$$\bar{h}_k^m(x) \sim \mathcal{CN}\left(0, \sum_{A:m\in A} \sum_{u\in \boldsymbol{\mathcal{T}}_k^{A}} P_r^{u_m} \ell(u,x)\right).$$

The independent phase rotations serve to de-correlate the channels[5] $\bar{h}_k^i(x)$ and $\bar{h}_k^j(x)$, $i, j \in \{1, 2\}$, $i \neq j$,. Without the phase rotations the channels would be correlated since some relays forward multiple symbols. This method can also be used to induce fluctuations in deterministic channels and improve performance, as in opportunistic beamforming [87]. Using (2.19) and letting $b = \{1, 2\}$, we have the following result.

Lemma 10 *Consider two independent users that wish to broadcast their messages using protocol PC, with equal power allocation. Under the continuum limits $N \to \infty$ and $P_r \to 0$, the rate of the aggregate channel gain $|\bar{h}_k^m(x)|^2$, an exponential RV, is*

$$\mu_k^m(x) = \left(\int_{\boldsymbol{\mathcal{S}}} \ell(u, x) \left(P_r \rho P_k^m(u) + \tfrac{P_r \rho}{2} P_k^b(u)\right) du\right)^{-1},$$

for $m = 1, 2$. Dropping the node index x, define the functions

$$\begin{aligned}
\theta_k(\mu^1, \mu^2) &= e^{-\tau N_o \mu_k^1}(1 - e^{-\tau N_o(\mu_k^2 + \tau \mu_k^1)})/(1 + \tau \mu_k^1/\mu_k^2), \\
\omega_k(\mu^1, \mu^2) &= e^{-\tau N_o \mu_k^1}/(1 + \tau \mu_k^1/\mu_k^2), \\
\gamma_k(\mu^1, \mu^2) &= \omega_k(\mu^1, \mu^2) - \theta_k(\mu^1, \mu^2), \\
\Gamma_k(\mu^1, \mu^2) &= (\omega_k(\mu^1, \mu^2) + \gamma_k(\mu^2, \mu^1)) \textstyle\prod_{i=1}^{k-1}(1 - \Gamma_i(\mu^1, \mu^2)).
\end{aligned}$$

Then, for $k > 0$, the per node level membership probabilities under the condition that $\tau > 1$ are

$$\begin{aligned}
P_{k+1}^1(x) &= \Gamma_k(\mu^1, \mu^2), \qquad P_{k+1}^2(x) = \Gamma_k(\mu^2, \mu^1), \\
P_{k+1}^b(x) &= (\gamma_k(\mu^1, \mu^2) + \gamma_k(\mu^2, \mu^1)) \\
&\qquad \textstyle\prod_{i=1}^{k-1}(1 - \omega_i(\mu^2, \mu^1) - \omega_i(\mu^1, \mu^2)).
\end{aligned}$$

Proof See Appendix A.5. ■

[5]This is easy to show because the channels $h(u, x)$ are ZMCSCG.

$(\mu_k^m(x))^{-1}$ has a nice interpretation as the power received at x from the cooperative transmission of s_m in slot k. While the expressions for the level probabilities do not provide insights directly, using them we can bound the decoding threshold τ for each user.

Lemma 11 *Protocol PC with equal power allocation can support the broadcast of the messages of two users provided the threshold τ is at most $\frac{2P_r\rho\log 3}{N_o d_o}$.*

Proof From Lemma 10,

$$\begin{aligned} P_{k+1}^1(x) &= \Gamma_k(\mu^1,\mu^2) \leq \omega_k(\mu^1,\mu^2) + \gamma_k(\mu^1,\mu^2) \\ &\leq e^{-\tau N_o \mu_k^1(x)} + e^{-\tau N_o \mu_k^2(x)}. \end{aligned} \tag{2.20}$$

Through similar manipulations we obtain identical upper bounds for $P_{k+1}^2(x)$ and $P_{k+1}^b(x)$. Using the definition of $\mu_k^m(x)$ from Lemma 10, $\mu_k^m(x) \geq$

$$(P_r\rho\left(\sup_{u\in\boldsymbol{\mathcal{S}}} P_k^m(u) + \tfrac{1}{2}\sup_{u\in\boldsymbol{\mathcal{S}}} P_k^b(u)\right)\int_{\boldsymbol{\mathcal{S}}}\ell(u,x)du)^{-1},$$

where $\int_{\boldsymbol{\mathcal{S}}}\ell(u,x)du = \frac{2}{d_o} - \frac{1}{d_o+x} - \frac{1}{d_o+L-x} \leq \frac{2}{d_o}$.

Let $g_{k+1}^i = \sup_{x\in\boldsymbol{\mathcal{S}}} P_{k+1}^i(x),\ i=1,2,b$ and let $\mathbf{g}_k = [g_k^1, g_k^2, g_k^b]^T$. The evolution of the level probabilities, in terms of the upper bounds, can now be written as a non-linear dynamical system

$$\mathbf{g}_{k+1} \leq (e^{\frac{-\tau N_o d_o}{2P_r\rho(g_k^1+g_k^b/2)}} + e^{\frac{-\tau N_o d_o}{2P_r\rho(g_k^2+g_k^b/2)}})\mathbf{1}, \tag{2.21}$$

where $\mathbf{1}$ is the all one vector. The solution to such non-linear systems is in general non-trivial.

We proceed by deriving a linear upper bound. Divide both sides of (2.21) by $g_k^1 + g_k^2 + g_k^b$. Using the Kuhn-Tucker conditions, we can show that the system below has a maximum at $\exp(-\tau N_o d_o / 2P_r\rho)$:

$$\max \frac{e^{\frac{-\tau N_o d_o}{2P_r\rho(g_k^1 + g_k^b/2)}} + e^{\frac{-\tau N_o d_o}{2P_r\rho(g_k^2 + g_k^b/2)}}}{g_k^1 + g_k^2 + g_k^b} \text{ s.t. } 0 \leq g_k^1, g_k^2, g_k^b \leq 1.$$

We can therefore write the dynamical system as

$$\mathbf{g}_{k+1} \leq \exp(-\tau N_o d_o / 2P_r\rho) B \ \mathbf{g}_k, \tag{2.22}$$

where B is a 3×3 all-one matrix. Equivalently, $\mathbf{g}_k \leq \exp(\frac{-k\tau N_o d_o}{2P_r\rho}) B^k \mathbf{g}_1$. The eigenvalues of B are $\lambda(B) = [3, 0, 0]^T$. Its clear that if $\exp(\frac{-\tau N_o d_o}{2P_r\rho})\lambda_{\max}(B) < 1$ then $\lim_{k\to\infty} \mathbf{g}_k = \mathbf{0}$, *i.e.*, all level probabilities vanish. Solving for τ such that this inequality is satisfied yields $\tau > 2P_r\rho \log 3 / N_o d_o$ as desired. ■

Naturally, the threshold cannot exceed this bound when $M > 2$ because interference will be greater. This leads to the following statement, analogous to Lemma 7, for the general case with M users.

Corollary 1 *The broadcast of M users using protocol PC cannot be successful if the decoding threshold τ is greater than $\frac{2P_r\rho \log 3}{N_o d_o}$ per user.*

2.4.2 Power Control with Side Information Protocol (PCSI)

The use of side-information intuitively provides performance gains over protocol PC. Below we provide the level probabilities for protocol PCSI with $M = 2$.

Lemma 12 *Let $P_k^i(x)$, $i = 1, 2, b$, denote the level probabilities for protocol PC. Then, the level probabilities for protocol PCSI, with equal power allocation, for $k > 1$ with $m, n \in \{1, 2\}$, $m \neq n$, are*

$$\begin{aligned}
\bar{P}_{k+1}^m(x) &= e^{-\tau N_o \mu_k^m(x)} \left(\sum_{j=1}^k \bar{P}_j^n(x)\right) \prod_{i=1}^k (1 - P_i^m(x)) \\
&\quad + P_{k+1}^m(x)\left(1 - \sum_{j=1}^k \bar{P}_j^n(x)\right), \\
\bar{P}_{k+1}^b(x) &= P_{k+1}^b(x), \quad \textit{where} \\
\mu_k^m(x) &= \left(\int_{\boldsymbol{\mathcal{S}}} \ell(u, x) \left(P_r \rho \bar{P}_k^m(u) + \tfrac{P_r \rho}{M} \bar{P}_k^b(u)\right) du\right)^{-1}
\end{aligned}$$

is the rate of the aggregate channel gain $|\bar{h}_k^m(x)|^2$ in the continuum limit. The level probabilities remain unchanged from protocol PC in $k = 1$.

Proof See Appendix A.6. ■

The level probabilities $\bar{P}_{k+1}^b(x)$ corresponding to the joint level set $\boldsymbol{\mathcal{T}}_{k+1}^b$ remain unchanged from protocol PC since $x \in \boldsymbol{\mathcal{T}}_{k+1}^b$ can only occur if x has no prior side-information. Verification of the expressions is provided in Section 2.5.

2.4.3 Single-User Receiver Protocol (SU)

Unlike protocols PC and PCSI which use SIC, level probabilities are simpler to calculate for SU because we have M unique decoding events ϵ_k^A, $A \in \boldsymbol{\mathcal{A}} = \{1, \ldots, M\}$, and therefore just M unique level sets. Therefore, we are able to analyze the protocol with M users.

Lemma 13 *Consider M independent users that wish to broadcast their messages using protocol SU. The rate of the aggregate channel gain $|\bar{h}_k^m(x)|^2$, an exponential*

RV, is

$$\mu_k^m(x) = \left(P_r \rho \int_{\boldsymbol{\mathcal{S}}} \ell(u,x) P_k^m(u) du\right)^{-1}. \tag{2.23}$$

Then, assuming $\tau > 1$, the level probabilities characterizing the evolution of the level sets are given by $P_{k+1}^m(x) =$

$$\prod_{\substack{i=1,\\ i\neq m}}^{M} \mu_k^i(x) \sum_{\substack{j=1,\\ j\neq m}}^{M} \frac{e^{-\tau N_0 \mu_k^m(x)} (\mu_k^m(x) + \frac{\mu_k^j(x)}{\tau})^{-1}}{\prod_{\substack{s=1,\\ s\neq j}}^{M} (\mu_k^s(x) - \mu_k^j(x))} \prod_{\ell=1}^{k} (1 - P_\ell^m(x)).$$

Proof The channel $\bar{h}_k^m(x) = \sum_{u \in \boldsymbol{\mathcal{T}}_k^m} \sqrt{P_r} h(u,x) e^{j\phi_u}$ is ZMCSG with variance $P_r \sum_{u \in \boldsymbol{\mathcal{T}}_k^m} \ell(u,x)$. Following [77], in the limit $N \to \infty$, $P_r \to 0$, the variance converges a.s. to $P_r \rho \int_{\boldsymbol{\mathcal{S}}} \ell(u,x) P_k^m(u) du$.

Denote the aggregate channel gain $|\bar{h}_k^m(x)|^2$ by Z_m. It is easy to see that Z_m has an exponential distribution with rate $\mu_k^m(x)$ as defined in the statement of the theorem. Z_m are mutually independent but non-identically distributed (*i.n.i.d.*) since the level sets $\boldsymbol{\mathcal{T}}_k^m$ must be disjoint but not identical.

The probability that node x belongs to level $\boldsymbol{\mathcal{T}}_{k+1}^m$ is the probability of the joint event that symbol s_m was received with the greatest power and $\mathrm{SINR}_{k+1}^m(x) \geq \tau$. Then, again using (2.19), we have $P_{k+1}^m(x) = P(x \in \boldsymbol{\mathcal{S}}_{k+1}^m) \prod_{\ell=1}^{k} P(x \in \bar{\boldsymbol{\mathcal{S}}}_\ell^m)$, where $P(x \in \bar{\boldsymbol{\mathcal{S}}}_\ell^m) = 1 - P_\ell^m(x)$ and

$$\begin{aligned} P(x \in \boldsymbol{\mathcal{S}}_{k+1}^m) &= P\left(\frac{Z_m}{N_o + \sum_{i\neq m} Z_i} \geq \tau \cap Z_m \geq \max_{i\neq m} Z_i\right) \\ &= \mathbb{E}_{Z_m} P\left(\sum_{i\neq m} Z_i \leq \frac{z_m - \tau N_o}{\tau} \cap \max_{i\neq m} Z_i \leq z_m\right) \\ &= \mathbb{E}_{Z_m} P\left(\sum_{i\neq m} Z_i \leq \frac{z_m - \tau N_o}{\tau}\right) \end{aligned} \tag{2.24}$$

where the last equality holds since the event $\sum_{i\neq m} Z_i \leq \frac{z_m - \tau N_o}{\tau}$ implies event $\max_{i\neq m} Z_i \leq z_m$ for $\tau > 1$. At this point, we note that the distribution of the sum $W = \sum_{i\neq m} Z_i$ of *inid* exponential random variables is known to be

$$f_W(w) = \left(\prod_{\substack{i=1,\\ i\neq m}}^{M} \mu_k^i(x)\right) \sum_{\substack{j=1,\\ j\neq m}}^{M} \frac{e^{-w\mu_k^j(x)}}{\prod_{\substack{s=1,\\ s\neq j}}^{M} (\mu_k^s(x) - \mu_k^j(x))}.$$

We use this distribution to compute the probability in (2.24) and then take the expectation with respect to Z_m, being mindful of the region of integration, $Z_m \in [\tau N_o, \infty]$, to obtain the final expression for the level probability. ■

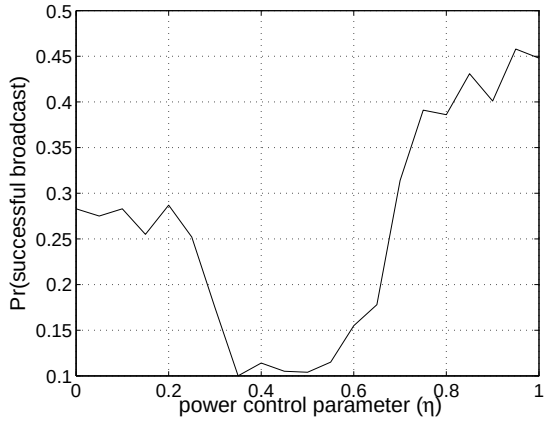
Next, using (2.24), we upper bound the level probability using the Markov inequality. For any $\theta > 0$, $P_{k+1}^m(x)$

$$\begin{aligned} &\leq \mathbb{E}_{Z_m}\left\{e^{\theta\frac{z_m-\tau N_o}{\tau}}\right\}\mathbb{E}\left\{e^{-\theta\sum_{i\neq m}Z_i}\right\}\prod_{\ell=1}^{k}(1-P_\ell^m(x)) \\ &= \frac{\mu_k^m(x)e^{-\tau N_o\mu_k^m(x)}}{\mu_k^m(x)-\theta/\tau}\prod_{\substack{i=1,\\ i\neq m}}^{M}\frac{\mu_k^i(x)}{\theta+\mu_k^i(x)}\prod_{\ell=1}^{k}(1-P_\ell^m(x)). \end{aligned}$$

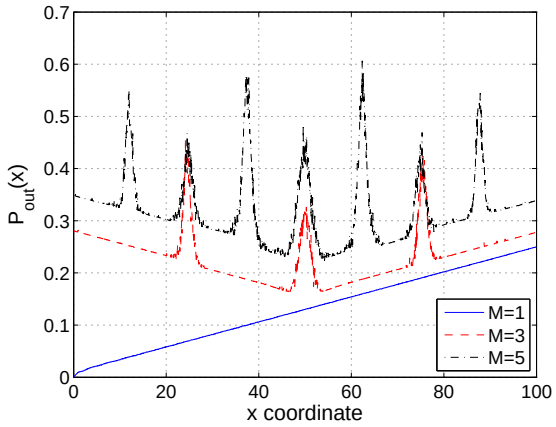
Recall, $\mu_k^i(x)$ is the reciprocal of the power received at x from the cooperative transmission of s_i. When the received power for s_i, for some i, is very large, $\mu_k^i(x)$ approaches 0, implying $P_{k+1}^m(x) \to 0$, $\forall m \neq i$. If we assume on the other hand that $\mu_k^i(x) > 0, \forall i$, it is clear that for M large enough, the middle product term vanishes $\forall\theta > 0$, forcing $P_{k+1}^m(x) \to 0$, $\forall m$, so not all messages can be broadcast.

The most likely scenario is that as M becomes large, the received power for many symbols s_i will become small since only a few relays will be able to decode s_i due to interference. So $\mu_k^i(x)$, for such i, will become large and $\frac{\mu_k^i(x)}{\theta+\mu_k^i(x)} \to 1$. In this environment with reduced interference, the propagation of a few flows $m \neq i$ will be able to continue and their broadcast will be successful. We call this the *capture effect*.

Thus, for M sufficiently large, $P_{k+1}^m(x) \to 0$, $\forall k > 1$, for at least some m. So single-user receivers as outlined in protocol SU do not permit effective multiplexing of the information of a large number of users because the system becomes interference limited.



(a)



(b)

Figure 2.4: (a) The optimal value of η is numerically found to be 0.95 for protocol PC (with fading), $M = 2$. (b) Outage probabilities for protocol PC (with fading).

2.5 Performance Analysis

We provide results supporting our analysis of the cooperative protocols under fading. The simulation setup is as follows. Consider a dense network of length $L = 100$ with $N = 1600$ nodes. Starting at coordinate 0, M users are located at intervals of length $\frac{L}{M-1}$. We assume that all user messages have a rate 2.5 bits/s/Hz and relay transmission power P_r.

Consider protocol PC under fading. For $M = 2$, Fig. 2.4(a) shows that the optimal value of the power control parameter is $\eta = 0.95$, verifying out intuition from Section 2.3.1. In fact, the worst broadcast success rate is observed when nodes that have decoded the messages of multiple users allocate equal power for forwarding each message ($\eta \approx 0.5$).

To understand the reliability of broadcast, define a *per node* outage probability $P_{out}(x)$. A node is in outage if it does not receive *all* the broadcast messages by the end of the session. Fig. 2.4(b) shows the outage probabilities for $M = 1, 3, 5$. Outages are most likely in the neighborhoods where two or more information flows cross - we observe M peaks. This is not surprising since nodes encounter the greatest interference in these regions. Outage probabilities for protocol PCSI (with fading) look very similar.

Now consider protocol PCSI with fading. We verify Lemma 12 for $M = 2$ users in Fig. 2.5. The two users are located at coordinates 0, and L and the two information flows originate from these points in $k = 0$. The analytical result from Lemma 12 is shown by smooth black curves. The level probabilities obtained via Monte Carlo trials are depicted by wiggly colored curves. Each color denotes a different time slot. We observe that the two sets of curves match, verifying the analytical expression. The slight discrepancy exists because we approximate an infinitely dense network with a finite number of nodes. Since probabilities $\bar{P}^m_{k+1}(x)$ are expressed in terms of the level probabilities $P^m_{k+1}(x)$ of protocol PC, this also verifies Lemma 10.

Finally, we compare the protocols SU and OIA, with fading, for $M = 3$ users located at 0, $L/2$, and L. We implemented OIA such that $D = 2$ dimensions are used per transmission. The level probabilities obtained numerically are plotted in

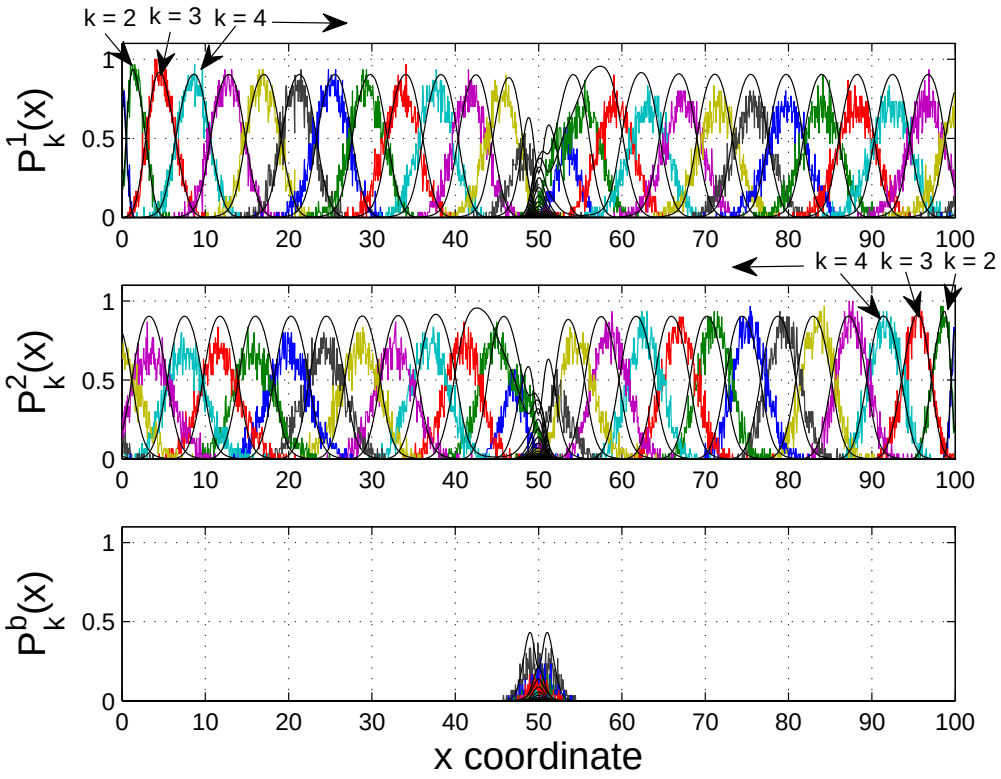
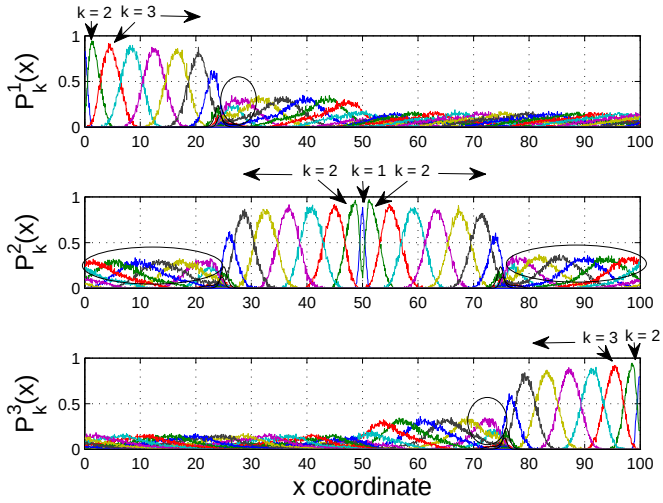


Figure 2.5: Verification of analytical level probability expressions for protocol PCSI ($M = 2$, fading channel). The expression from Lemma 12 (shown in smooth black curve) is superimposed on the Monte Carlo level probabilities (shown in squiggly colored curves).

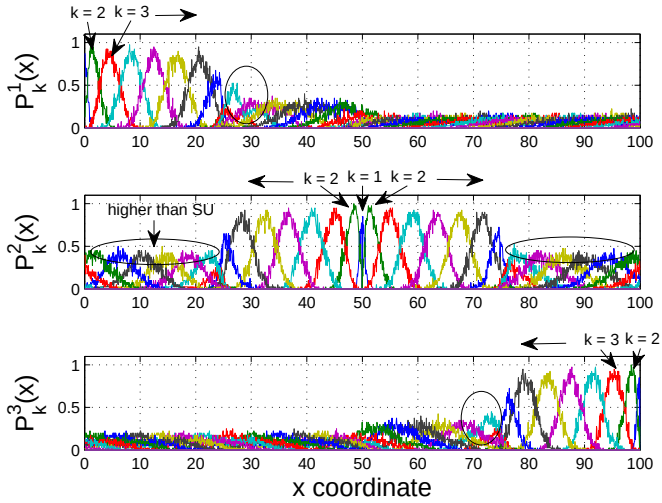
Figs. 2.6(a)-2.6(b). OIA displays higher level probabilities as marked by the black ellipses, implying greater chances of successful broadcast. As M becomes large, D could be increased in parallel (but a a slower rate to maintain $D \ll M$) to improve the likelihood of successful broadcast.

2.5.1 Comparison with Flooding Algorithms

We now compare the cooperative protocols to non-cooperative flooding algorithms. Probabilistic flooding [72] is one class of algorithms that optimizes flooding. The



(a)



(b)

Figure 2.6: Numerical level probabilities for protocol SU (a) and for protocol OIA (b) are shown for $M = 3$ users (both with fading). Timeslots are color coded and indicated with text. The black ellipses indicate the areas where OIA exhibits higher level probabilities.

idea is to minimize channel contention and message redundancy by requiring each node to forward the message with a certain forwarding probability p. These algorithms degenerate to pure flooding algorithms when $p = 1$. Furthermore, both these algorithms are distributed and very simple in terms of the resources required

of the nodes, as opposed to say [69, 90, 79, 33]. They are, therefore, similar to the cooperative protocols in terms of overheads. To benchmark the proposed cooperative broadcast protocols we compare their latency with the probabilistic and pure flooding algorithms.

Message passing in non-cooperative flooding algorithms is asynchronous while the cooperative protocols assume a time-slotted network. To have a meaningful comparison, we simulated a centralized implementation of the flooding algorithms in a synchronous time-slotted network. The duration of each time slot is the actual message duration plus the propagation and processing delays such that the nodes can complete all forwarding or reception and decoding operations in one time slot. The time slot duration for the cooperative protocols is the same.

Two nodes u and x are said to be connected if $|u - x| \leq R$, where R is the transmission range and is picked in the following manner for a fair comparison. We assume that all user messages have a rate 2.5 bits/s/Hz and relay transmission power P_r, in both the flooding algorithms and the cooperative protocols. Messages transmitted by node u are received and decoded by node x in an error-free manner provided that $\log(1+\text{SNR}(u,x)) \geq 2.5$ bits/s/Hz and there are no packet collisions. Using the usual definition $\text{SNR}(u,x) = P_r \ell(u,x)/N_o$, with path loss exponent $\beta = 2$, this condition is equivalent to $R \leq \sqrt{\frac{P_r}{\tau N_o}} - d_o$. Similar results are obtained for $\beta > 2$.

Nodes are half-duplex and are assigned indices. If a node x receives user $m's$ message in slot k successfully and decides to forward it, then it does so in slot $k+1$. CSMA/CD MAC is used, *i.e.*, the network uses both carrier sensing and collision detection. If neighbors of x also attempt to transmit messages in slot $k+1$ then all but one of the transmitting nodes detect that the medium is busy and back

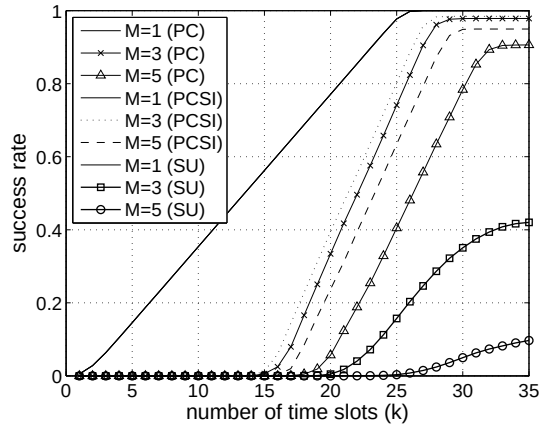
off. Since our implementation is centralized, the node with the smallest index is allowed to complete its transmission. Similar results are observed when a random node is allowed to win the contention instead.

Define the success rate to be the fraction of nodes that receive the messages broadcast by *all* users. We fix success rate > 0.9 for both flooding algorithms and the cooperative protocols (except protocol SU) in our simulations.

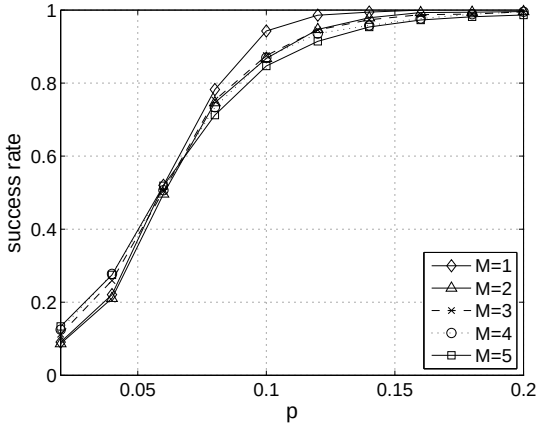
Fig. 2.7(a) highlights that because the cooperative protocols operate purely in the physical layer and there are no retransmissions, 100% success rates in a fading channel are not guaranteed. In contrast, probabilistic flooding utilizes CSMA/CD, so one can achieve perfect success rates provided the p is chosen appropriately.

Fig. 2.7(a) shows the success rates versus M for cooperative protocols PC and PCSI under the fading channel, using equal power allocation. Notably, protocol PCSI yields better success rates as M increases. This demonstrates the gains extracted by using side-information. Fig. 2.7(a) also shows that for the same parameters as the other two protocols, the success rates for protocol SU with fading are very low. In other words, as M grows, the broadcast has a low probability of success if nodes use single-user receivers. The performance for $M = 1$ is identical in all protocols, as expected. Therefore, if the application only requires the broadcast of a single user's message then the simplest protocol (SU) should be used.

Fig. 2.7(b) shows the success rate achieved by probabilistic flooding with respect to the forwarding probability p for M users. We see a success rate > 0.9, for $p \geq 0.16$, $\forall M$. There exists an optimal p which minimizes the latency for each M. We found that $p = 0.26$, (with success rate > 0.95) is optimal for $M \leq 5$. We use this value in our simulations.



(a)



(b)

Figure 2.7: (*a*) M users wish to flood the network. The plot shows the success rates achieved by cooperative protocols SU, PC, and PCSI (with equal power allocation where applicable, all with fading) for different values of M. The performance for $M = 1$ is identical in all protocols. (*b*) The plot shows the success rate as a function of the forwarding probability p for each M for probabilistic flooding.

Using Monte Carlo simulations we determined the average latency (in terms of the number of slots) required by the probabilistic flooding algorithm to flood the network. This is shown in Fig. 2.8. The delay performance of pure flooding is also provided for reference. As the number of independent users broadcasting their messages increases, the average latency scales as $\Omega(M)$, using standard big-oh

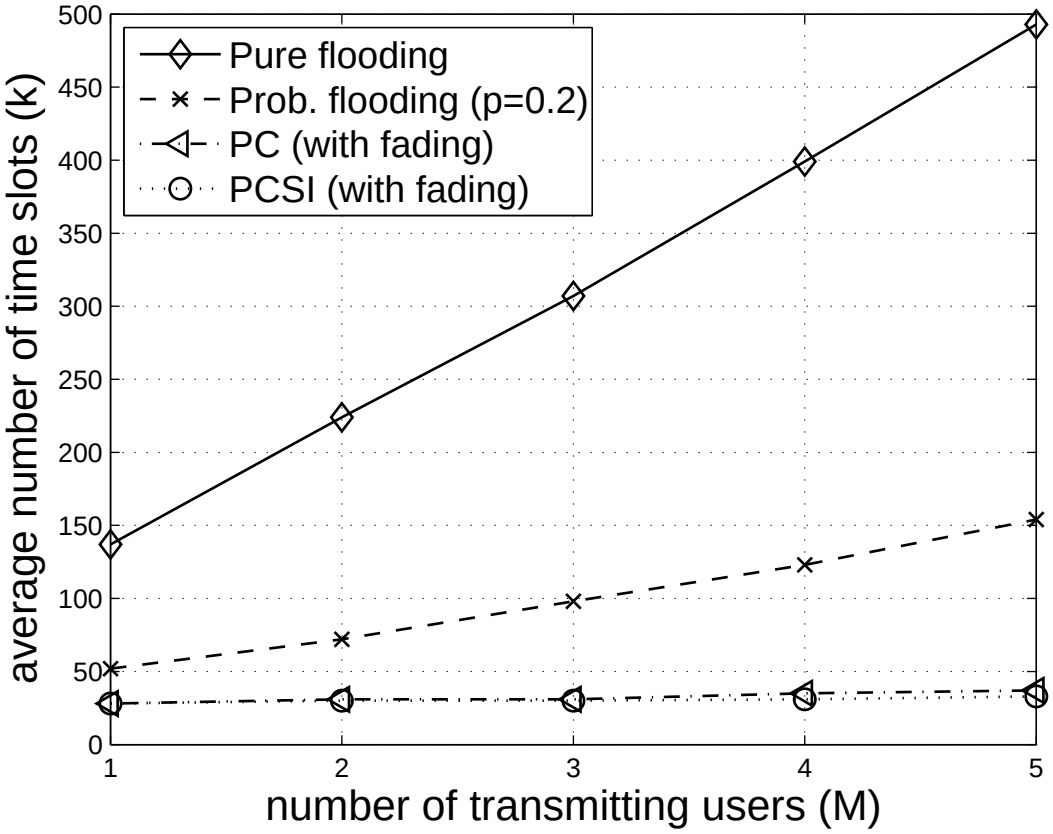


Figure 2.8: The plot shows the average number of time slots required to flood the network at a success rate > 0.9 for the two flooding algorithms and the cooperative protocols. The cooperative protocols achieve lower latency.

notation.

Cooperative protocols, with fading, outperform probabilistic flooding for $M = 1$ and the latency grows negligibly as M increases. This hints at a very favorable end-to-end latency scaling law for multiplexing information using the cooperative protocols. In particular, Fig. 2.8 shows the latency for protocols PC and PCSI for the fading channel model. For $M = 5$, protocol PCSI (with side-information) shows an improvement of about 10% over protocol PC in terms of latency (33 vs

37 slots), but more significantly, for the same parameters its success rate is 46% higher (see Fig. 2.7(a)). Furthermore, for $M = 5$, protocol PCSI decreases latency by 84% compared to probabilistic flooding. Protocol SU fails to satisfy the desired success rate for the same parameters, so its not shown.

The simulation results show that protocol PCSI displays the best broadcast latency. The improvement in performance by using side-information comes at the cost of increased memory requirements for the nodes since they must buffer each message they have decoded. In our simulations, we allowed each relay to store up to $M - 1$ unique signals from other users that it has decoded previously. Given M users in the network, the required memory per relay is upper bounded by $M - 1$ units. As the maximum memory allowed per relay is reduced, the average number of time slots required for successful broadcast by protocol PCSI (Fig. 2.8) will increase and trend toward the performance obtained by protocol PC.

CHAPTER 3

COOPERATIVE BROADCAST OF SINGLE SOURCE: ERROR PROPAGATION

In the previous chapter we considered cooperative protocols for broadcasting the content of multiple sources simultaneously. We assumed, however, that relays made no errors during decoding. In this chapter, we consider what happens if we relax that assumption. For clarity, we now focus on the broadcast of the content of a single source for obtaining insight into the error propagation. There are a few other differences in the model employed as specified below.

3.1 System Model

Consider a source node that wants to broadcast its message to a certain predetermined region of the network. Assume that N nodes are deployed with a random uniform distribution in this region of unit area. For simplicity, we will assume that the region is a strip of width W and length L ($WL = 1$) and that the source node located at one edge of the strip transmits with power P_s. We also assume that the source and relay data are BPSK modulated. The network region is divided into consecutive areas L_k, $k = 1, \ldots, K$. For the simple strip model consider:

$$L_k = \{(k-1)L/K \leq x < kL/K, |y| \leq W/2\}, k = 1, \ldots, K$$

while $L_0 = (0, 0)$ is only one point. All nodes that fall within L_k form a *level* k set

$$\mathcal{S}_k = \{i : (x_i, y_i) \in L_k,\ i = 1, \ldots, N\}, \quad k = 1, \ldots, K.$$

The source is the only node in level zero *i.e.,* $\mathcal{S}_0 = \{i = 0, (x_0, y_0) = (0, 0)\}$. For our analysis we consider two cases: a) the case where the cooperative nodes in each

level use *orthogonal channels*, such as Orthogonal Space Time Codes (OSTC), or simply TDMA/FDMA or CDMA scheduling; b) the case where each level uses only one signal dimension and hence the transmissions are *non-orthogonal*.

Clearly, in the first case we need a bandwidth expansion proportional to the number of nodes in the network, because we need $|\mathcal{S}_k|$ channels per level and each level transmits over a different time slot. Instead, in the second case, we only require a bandwidth expansion on the order of the number of levels K, which is comparable to what a non cooperative multi-hop system requires. In our model we also assume that all nodes in each level $\mathcal{S}_k$ transmit synchronously at the symbol level. Since the events occur in a chain, synchrony at the symbol level can be enforced if the transmission is narrow-band and by having each level estimate the time of arrival of the previous level packet and retransmit within a deadline from its arrival time. Small scale asynchronism is incorporated in our model because we assume the presence of fading. In either case, we assume that the nodes can detect whether the previous level transmitted a message and what the link gain of each channel is from a preamble[1].

3.2 Orthogonal Transmissions

3.2.1 Random Finite Network

We assume that the orthogonal channels experience random flat fading and the link coefficient is $\beta_{ji} \triangleq \alpha_{ji} e^{j\theta_{ji}} \sqrt{P_r \ell(d_{ji})}$ where P_r is the power of the relay transmissions. The fading is characterized by deterministic large scale fading factors

[1]Depending on the relay powers, this is actually not always possible; for further details see [51].

$\ell(d_{ji})$ which are functions of the internode distance (d_{ji}) only, envelopes α_{ji} that may or may not be independent and random phases θ_{ji} that are i.i.d. uniform in $[0, 2\pi]$. The latter assumption is justified by the fact that nodes have independent oscillators and separate RF front-ends. The network is pre-partitioned as described in the previous section and we assume that the source transmits a BPSK message with power P_s. Then, for each symbol b transmitted by the source, the discrete time baseband complex equivalent model for the corresponding received symbol at the j^{th} node in $\mathcal{S}_1$ is

$$r_j^{(1)} = b\alpha_{j0}e^{j\theta_{j0}}\sqrt{P_s\ell(d_{j0})} + w_j, \tag{3.1}$$

where α_{j0} is the small scale fading factor at node j, θ_{j0} is the carrier phase shift, w_j is a circularly symmetric complex Gaussian noise sample with zero mean and variance N_o, d_{j0} is the distance between the source and the j^{th} node, and $\ell(d_{j0})$ is the path-loss attenuation function which we assume to be deterministic (*e.g.* for the free-space model it is $1/d_{j0}^2$). Because we assume that the channel gain has been estimated perfectly by all nodes in the level, the nodes in $\mathcal{S}_1$ can use a coherent detector and the Bit Error Rate (BER) at a node j in $\mathcal{S}_1$ is given by

$$P_e^{(1)} = Q\left(\sqrt{\frac{2P_s}{N_o}\left|\alpha_{j0}\right|^2\ell(d_{j0})}\right). \tag{3.2}$$

where the Q-function is defined as $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$. As specified by the cooperative transmission protocol in Section 3.1, the relay nodes in $\mathcal{S}_1$ will transmit their decoded bits over orthogonal channels. Subsequently, the relay nodes in all other levels will do the same. But not all relay nodes will decode the bit they received correctly, therefore errors will be introduced into the bit retransmission flow at the physical layer. Let the orthogonal channel between the relay nodes $i \in \mathcal{S}_k$ and $j \in \mathcal{S}_{(k+1)}$ have a channel gain $\beta_{ji} \triangleq \alpha_{ji}e^{j\theta_{ji}}\sqrt{P_r\ell(d_{ji})}$. The received signal at the jth node in level $\mathcal{S}_{(k+1)}$ can be modelled as a vector of symbols received

over individual orthogonal channels corresponding to the retransmission of bit b from nodes in level $\mathcal{S}_k$. Thus the received signal can be expressed as:

$$\boldsymbol{r}_j^{(k+1)} = b\left(\boldsymbol{\beta}_j \odot \boldsymbol{\epsilon}^{(k)}\right) - 2b\left(\boldsymbol{\beta}_j \odot \boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)}\right) + \boldsymbol{w}_j \tag{3.3}$$

where the operator $\odot$ represents element-wise multiplication, $\boldsymbol{\beta}_j$ is a $N \times 1$ vector (with i^{th} element β_{ji}) specifying the link gains of all the orthogonal channels in the network, $\{\boldsymbol{w}_j\}_i \sim \mathcal{CN}(0, N_o)$, and $\boldsymbol{\epsilon}^{(k)}$ and $\boldsymbol{e}^{(k)}$ are $N \times 1$ vectors whose i^{th} element is specified by

$$\epsilon_i^{(k)} = \begin{cases} 1 & \text{if the } i^{th} \text{ node} \in \mathcal{S}_k, \\ 0 & \text{otherwise} \end{cases}$$

$$e_i^{(k)} = \begin{cases} 1 & \text{if the } i^{th} \text{ node in } \mathcal{S}_k \text{ makes an error} \\ 0 & \text{otherwise.} \end{cases}$$

For a given network deployment $\boldsymbol{\epsilon}^{(k)}$ is known a priori if the node clustering is known. However, $\boldsymbol{\epsilon}^{(k)}$ is random since the deployment is random. As mentioned in Section 3.1, we assume that node j knows the combined vector $\boldsymbol{\beta}_j \odot \boldsymbol{\epsilon}^{(k)}$ without error.

Vector $\boldsymbol{e}^{(k)}$, which indicates the nodes that make an error in $\mathcal{S}_k$, is random and unknown to the next level. Information concerning the statistics of the error vector $\boldsymbol{e}^{(k)}$ should be incorporated in the construction of the Maximum-Likelihood detectors at the receivers. However, for ease of analysis we consider a suboptimal receiver that combines the observations in $\boldsymbol{r}_j^{(k+1)}$ using a Maximal-Ratio Combining (MRC) receiver. Hence, our analysis of the error propagation is pessimistic. The MRC detector rule is [68]:

$$\Re\left\{\left(\boldsymbol{\beta}_j \odot \boldsymbol{\epsilon}^{(k)}\right)^H \cdot \boldsymbol{r}_j^{(k+1)}\right\} \gtrless 0. \tag{3.4}$$

Let us define

$$z_j^{(k+1)} \triangleq \left\| \boldsymbol{\beta}_j \odot \boldsymbol{\epsilon}^{(k)} \right\|^2 = \sum_{i=1}^{N} |\beta_{ji}|^2 \epsilon_i^{(k)}, \tag{3.5}$$

$$v_j^{(k+1)} \triangleq \left\| \boldsymbol{\beta}_j \odot \boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)} \right\|^2 = \sum_{i=1}^{N} |\beta_{ji}|^2 \epsilon_i^{(k)} e_i^{(k)}. \tag{3.6}$$

Then the bit error rate of the j^{th} node in level L_{k+1} is

$$P\left(e_j^{(k+1)} = 1 | \mathcal{A}\right) = Q\left(\frac{z_j^{(k+1)} - 2v_j^{(k+1)}}{\sqrt{z_j^{(k+1)} \frac{N_o}{2}}}\right) \tag{3.7}$$

where $\mathcal{A} = \boldsymbol{\beta}_j$, $\boldsymbol{\epsilon}^{(k)}$, and $\boldsymbol{e}^{(k)}$ for the previous k.

These are the basic equations needed to analyze how the error propagates through the network; which essentially corresponds to evaluating the statistics of the vector $\boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)}$. For different values of k, the vector $(\boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)})$ is a Markov-chain whose statistics are cumbersome to analyze. Hence, to understand the behavior of the network we use a combination of asymptotic results that are valid when taking the limit for $N \to \infty$ while decreasing the transmit power $P_r \to 0$.

3.2.2 Continuum Network

In the continuum model we are interested in the behavior of *high density* networks with *constant sum-power*, i.e. the number of nodes, N, goes to infinity while the total relay power, $P_r N$, is fixed. This means that the number of nodes cooperating in each level of the cooperative transmission increases to infinity, however, the power density in each level remains finite.

We make use of the following assumptions in our analysis: (a) The positions of all nodes (x_i, y_i) are independent and identically distributed (i.i.d.), (b) The small

scale fading coefficients α_{ji} have unit variance the phases θ_{ji} are i.i.d. uniform $[0, 2\pi]$, (c) d_{ji}, α_{ji}, and θ_{ji} are statistically independent $\forall j, i$, and (d) $\boldsymbol{w}_j$ is AWGN with zero mean and variance N_o.

In force of all the assumptions listed above, $\boldsymbol{\epsilon}^{(k)}$ and $\boldsymbol{e}^{(k)}$ are statistically independent. Under the above stated conditions the following lemma holds:

Lemma 14 *Assume that the network has unit area. In the asymptote, as $N\to\infty$ and $P_r\to 0$, by fixing $\lim_{N\to\infty}\lim_{P_r\to 0} P_r N = \bar{P}_r$ for every level $k > 1$ we get the following relations:*

The probability that the j^{th} node, located at coordinates (x, y) belongs to level L_k is such that

$$\lim_{N\to\infty} E\{\epsilon_j^{(k)}\} = \pi^{(k)}(x, y) = \begin{cases} 1 & (x, y) \in L_k, \\ 0 & \textit{otherwise.} \end{cases} \tag{3.8}$$

With probability 1, $z_j^{(k+1)}$ converges to a deterministic value:

$$\begin{aligned} \lim_{\substack{N\to\infty \\ P_r\to 0}} z_j^{(k+1)} &= \xi^{(k+1)}(x, y) \\ &= \bar{P}_r \iint \ell(x-u, y-v)\pi^{(k)}(x, y)dudv \\ &= \bar{P}_r \iint_{L_k} \ell(x-u, y-v)dudv \end{aligned} \tag{3.9}$$

where the integration in the middle step is over the entire network area and $\ell(x-u, y-v)$ is the path-loss attenuation function between the nodes at coordinates (x, y) in level $\mathcal{S}_{k+1}$ and (u, v) in level $\mathcal{S}_k$. By letting $\lim_{\substack{N\to\infty \\ P_r\to 0}} E\{e_i^{(k)}\} = \psi^{(k)}(x, y)$, the coefficient $v_j^{(k+1)}$ tends to the following integral with probability 1:

$$\begin{aligned} \lim_{\substack{N\to\infty \\ P_r\to 0}} v_j^{(k+1)} &= \nu^{(k+1)}(x, y) \\ &= \bar{P}_r \iint_{L_k} \ell(x-u, y-v)\psi^{(k)}(u, v)dudv, \end{aligned} \tag{3.10}$$

and for all $k \geq 2$:

$$\psi^{(k)}(x,y) = Q\left(\frac{\xi^{(k)}(x,y) - 2\nu^{(k)}(x,y)}{\sqrt{\frac{N_o}{2}\xi^{(k)}(x,y)}}\right). \tag{3.11}$$

The set of equations above can be solved starting for $k = 1$ by using:

$$\psi^{(1)}(x,y) = E_\gamma\left\{Q\left(\sqrt{\frac{2P_s}{N_0}\gamma\ell(x,y)}\right)\right\} \tag{3.12}$$

where γ is the square of the fading envelope $|\alpha_{ji}|e^{j\theta_{ji}}$.

Proof For a given N and P_r, denote the mean of $z_j^{(k+1)}$ by $M_N^{(k+1)}$. The following holds for $k \geq 1$:

$$\begin{aligned} M_N^{(k+1)} &\triangleq \mathbb{E}_\alpha\left\{z_j^{(k+1)}\right\} = \sum_{i=1}^{N}\mathbb{E}_\alpha\left\{|\beta_{ji}|^2\,\epsilon_i^{(k)}\right\} \\ &= \frac{\bar{P}_r}{N}\sum_{i=1}^{N}\ell(d_{ji})\mathbb{E}\left\{\epsilon_i^{(k)}\right\} \end{aligned} \tag{3.13}$$

Because the nodes are deployed randomly under a uniform distribution (see Theorem 1 in [77]) we get

$$\lim_{N\to\infty} M_N^{(k+1)} = \bar{P}_r\iint_{L_k}\ell(x-u,y-v)dudv \tag{3.14}$$

By the Law of Large Numbers (LLN), (3.9) follows.

Now let:

$$\psi^{(k)}(x_i,y_i) = \mathbb{E}\left\{e^{(k)}(x_i,y_i)\right\} \tag{3.15}$$

Arguing similarly for $v_j^{(k+1)}$, $\bar{M}_N^{(k+1)} \triangleq \mathbb{E}_\alpha\{v_j^{(k+1)}\}$ is

$$\bar{M}_N^{(k+1)} = \frac{\bar{P}_r}{N}\sum_{i=1}^{N}\ell(x_j-u_i,y_j-v_i)\mathbb{E}\{\epsilon_i^{(k)}\}\mathbb{E}\{e_i^{(k)}\}$$

$$\lim_{N\to\infty}\bar{M}_N^{(k+1)} = \bar{P}_r\iint_{L_k}\ell(x-u,y-v)\psi^{(k)}(u,v)dudv.$$

By the LLN this leads to (3.10).

Because $Q(x)$ is a continuous function, the two results above imply that for $k \geq 2$ equation (3.11) is valid. ■

Hence, what Lemma 14 states is that the expected probability of error for each value of the coordinates (x, y) is governed by a set of non-linear difference-integral equations, *i.e.,* (3.9), (3.10), and (3.11). The functions $\psi^{(k)}(x, y)$ that solve the equations in each region L_k represent the network error *dynamics*. By definition, $0 \leq \psi^{(k)}(x, y) \leq 0.5$. To evaluate the exact behavior of the error propagation across the levels in the asymptotic regime one must numerically evaluate the solutions for these equations for every value of $k = 1, ..., K$ (see Sec. 3.4).

Next, we obtain analytical insight on the error performance by finding bounds for $\psi^{(k)}(x, y)$ for every $k = 1, \ldots, K$. Catastrophic error propagation can be avoided if it is possible to choose L_k and/or $\bar{P}_r$ so that $\max \psi^{(k)}(x, y) \leq \lambda < 0.5$. The following lemma shows when and how this is possible:

Corollary 2 *Assuming that all L_k cover an equal area, let:*

$$\alpha = \min_{(x,y)} \sqrt{\frac{2\bar{P}_r}{N_o} \iint_{L_k} lemma.ortho(x-u, y-v)dudv} \quad \forall k = 1, \ldots, K. \tag{3.16}$$

Under the same assumptions that lead to Lemma 1, the worst error probability reaches one of the fixed points i.e., $\max\left(\psi^{(\infty)}(x, y)\right) \to \lambda$*, that is a solution of:*

$$\lambda = Q(\alpha(1 - 2\lambda)) \tag{3.17}$$

A trivial fixed point that can be achieved $\forall$ α is $\lambda = 0.5$. It is the only fixed point for $0 \leq \alpha \leq \sqrt{\frac{\pi}{2}}$ and it means that for α in the above range, catastrophic

error propagation will occur with probability one. For $\alpha > \sqrt{\frac{\pi}{2}}$ *another fixed point exists*[2]*:*

$$Q(\alpha) < \lambda < \frac{1}{2} - \frac{1}{2}\sqrt{-\frac{2}{\alpha^2}\ln\sqrt{\frac{\pi}{2\alpha^2}}}.$$

If $\psi^{(1)}(x,y)$ *is smaller than* $\frac{1}{2} - \frac{1}{2}\sqrt{-\frac{2}{\alpha^2}\ln\sqrt{\frac{\pi}{2\alpha^2}}}$*, the worst error performance will be bounded by the fix point of (3.17).*

Proof Substituting the expressions for $\xi^{(k)}(x,y)$ and $\nu^{(k)}(x,y)$ into (3.11)

it is not difficult to see that

$$\max\left(\psi^{(k)}(x,y)\right) \leq Q\left(\alpha\left[1 - 2\max(\psi^{(k-1)}(x,y))\right]\right) \tag{3.18}$$

If we replace $\max\psi^{(k-1)}(x,y)$ and $\max\psi^{(k)}(x,y)$ with λ_{k-1} and λ_k respectively in (3.18), we get:

$$\lambda_k \leq Q(\alpha(1 - 2\lambda_{k-1})) \tag{3.19}$$

which implies that λ will tend to be limited to at most one of the exact solutions of the equation (3.19).

When $\lambda = 0.5$ the equality is met no matter what α is. Because $0 \leq BER \leq 0.5$ we only need to analyze if another solution exists in this range. The function $Q(\alpha(1-2\lambda))$ is monotonically increasing in this range and is concave because:

$$\frac{dQ(\alpha(1-2\lambda))}{d\lambda} = \sqrt{\frac{2\alpha^2}{\pi}}e^{-\frac{\alpha^2(1-2\lambda)^2}{2}} \geq 0.$$

For $\lambda = 0$ the function $Q(\alpha) > 0$ for any finite α. Hence, $Q(\alpha(1-2\lambda))$ and the straight line have another intersection point for $\lambda < 0.5$ if α is such that the point

[2]Note that for $\alpha \gg 1$ $\lambda \approx Q(\alpha)$ is a fixed point.

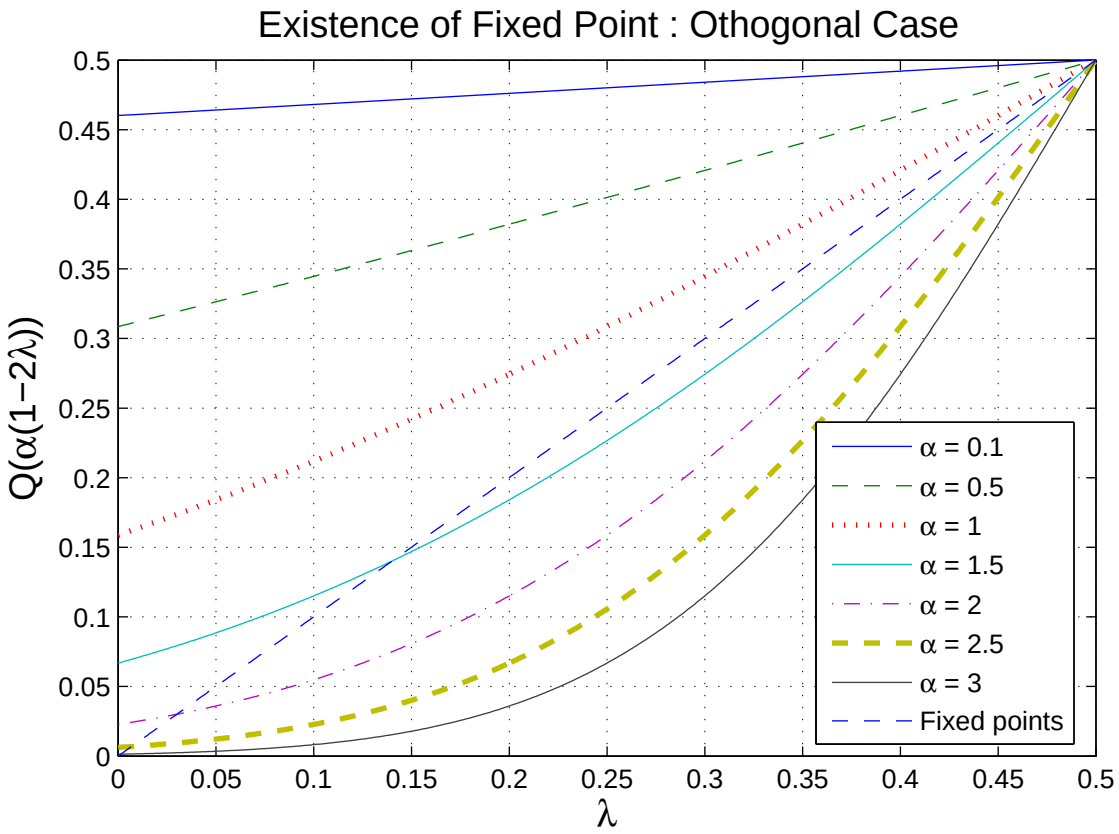


Figure 3.1: λ has a fixed point for some values of $Q(\alpha)$

where $Q(\alpha(1-2\lambda))$ has tangent 1 is less than $\lambda = 0.5$ (see Figure 1). The value of $\alpha_\star$ for which the phase transition happens is:

$$\left.\frac{dQ(\alpha_\star(1-2\lambda))}{d\lambda}\right|_{\lambda=0.5} = \sqrt{\frac{2\alpha_\star^2}{\pi}} = 1 \quad \Rightarrow \quad \alpha_\star = \sqrt{\frac{\pi}{2}}.$$

For all $\alpha > \alpha_\star = \sqrt{\frac{\pi}{2}}$, two intersection points exist: one is still at $\lambda = 0.5$ and the other one is for a $\lambda \geq Q(\alpha)$.

The second intersection point also has to be below the point $\lambda_\star$ where $\frac{dQ(\alpha(1-2\lambda_\star))}{d\lambda} = 1$ where for $\alpha > \alpha_\star$ it holds that $Q(\alpha(1-2\lambda_\star)) < \lambda_\star$. Thus for $\alpha > \sqrt{\frac{\pi}{2}}$, $Q(\alpha) \leq \lambda \leq \lambda_\star$. The value of $\lambda_\star$ is calculated easily, by solving $\sqrt{2\alpha^2/\pi}e^{-\frac{\alpha^2(1-2\lambda_\star)^2}{2}} = 1$ which gives $\lambda_\star = 0.5 - 0.5\sqrt{-\frac{2}{\alpha^2}\ln\sqrt{\frac{\pi}{2\alpha^2}}}$. ■

3.2.3 Comparison with a Non-Cooperative Network

For $\alpha > \sqrt{\frac{\pi}{2}}$, Lemma 14 shows that the BER over each level of the cooperative network is going to be bounded. Suppose that the cooperative physical layer is replaced by a single node at the edge of each level that can transmit at the total power of an entire level. Denoting the best point at each level to forward the message by (x_k, y_k), the average BER from level to level is $P_{k,k-1} = \mathbb{E}_{\gamma_k}\left\{Q\left(\frac{2\bar{P}_r|L_k|}{N_o}\gamma_k lemma.ortho(x_k - x_{k-1}, y_k - y_{k-1})\right)\right\}$ where γ_k is the square of the fading envelope of the "hop." Therefore, the BER for the multi-hop network is

$$BER = 1 - \prod_{k=1}^{\mathcal{K}} (1 - P_{k,k-1}) \approx \mathcal{O}(\mathcal{K}) \tag{3.20}$$

because $BER \geq 1 - (1 - \min_k (P_{k,k-1}))^{\mathcal{K}} = \mathcal{O}(\mathcal{K})$ if $\min_k(P_{k,k-1}) \ll 1$. This indicates that eventually the error will grow to 0.5. The comparison made above is not entirely fair because the cooperative network we analyzed uses infinite bandwidth whereas the non-cooperative multi-hop network does not. Below we show that this asymptotic bandwidth expansion is unnecessary. The low achievable bounds on the error dynamics we have found above for the cooperative scheme using orthogonal channels are shared by a scheme that utilizes an amount of bandwidth comparable to that of the multi-hop scheme.

3.3 Non-orthogonal Transmissions

3.3.1 Random Finite Network

The problem setup is the same as that described for orthogonal transmission in Section 3.2.1 except that the nodes in the same level are no longer assigned orthogonal channels for their transmissions. Hence, for every transmitted binary symbol there will be only one received coefficient per level. Using similar notation and the same definitions for $\boldsymbol{\beta}_j$, $\boldsymbol{\epsilon}^{(k)}$ and $\boldsymbol{e}^{(k)}$ as in the orthogonal transmission case, the sample received by the j^{th} node in level L_{k+1} which corresponds to a certain binary symbol b is:

$$r_j^{(k+1)} = b\left(\boldsymbol{\beta}_j^T \cdot \boldsymbol{\epsilon}^{(k)}\right) - 2b\boldsymbol{\beta}_j^T \cdot \left(\boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)}\right) + w_j^{(k+1)}. \tag{3.21}$$

Note that $r_j^{(k+1)}$ is a now a scalar quantity and $w_j^{(k+1)}$ is the sole noise sample that distorts the reception of binary symbol b over the shared cooperative channel. Let us define

$$z_j^{(k+1)} \triangleq \boldsymbol{\beta}_j^T \cdot \boldsymbol{\epsilon}^{(k)} \quad \text{and} \quad v_j^{(k+1)} \triangleq \boldsymbol{\beta}_j^T \cdot \left(\boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)}\right).$$

Again, assume that during the training phase the receivers estimate $z_j^{(k+1)}$ accurately. The coherent detector uses the following decision rule

$$\Re\left\{\left(z_j^{(k+1)}\right)^* \cdot r_j^{(k+1)}\right\} \gtrless 0. \tag{3.22}$$

The bit error rate of the j^{th} node in level L_{k+1} for $k \geq 2$ is

$$P\left(e_j^{(k+1)} = 1|\mathcal{A}\right) = Q\left(\frac{\left|z_j^{(k+1)}\right|^2 - 2\Re\left\{z_j^{*(k+1)} \cdot v_j^{(k+1)}\right\}}{\sqrt{\frac{N_o}{2}\left|z_j^{(k+1)}\right|^2}}\right) \tag{3.23}$$

where $\mathcal{A} = \boldsymbol{\beta}_j$, $\boldsymbol{\epsilon}^{(k)}$, and $\boldsymbol{e}^{(k)}$ for the previous k.

Once again, rather than analyzing the Markov-chain $\left(\boldsymbol{\epsilon}^{(k)} \odot \boldsymbol{e}^{(k)}\right)$ for a finite network we analyze its asymptote in the following section.

3.3.2 Continuum Network

The following lemma holds for a network of unit area.

Lemma 15 *In the asymptote, as $N\to\infty$ and $P_r\to 0$, by fixing $\lim_{N\to\infty}\lim_{P_r\to 0} P_r N = \bar{P}_r$ we get the following relations - Let $\lim_{\substack{N\to\infty \\ P_r\to 0}} \mathbb{E}\left\{e_j^{(k)}\right\} = \psi^{(k)}(x,y)$. The probability that the j^{th} node belongs to level L_k is:*

$$\lim_{N\to\infty} E\{\epsilon_j^{(k)}\} = \pi^{(k)}(x,y) = \begin{cases} 1 & (x,y) \in L_k, \\ 0 & otherwise \end{cases} \tag{3.24}$$

The coefficients $\lim_{N\to\infty}\lim_{P_r\to 0}(z_j^{(k+1)}, v_j^{(k+1)})$ converge in distribution to a zero mean jointly Gaussian circularly symmetric random variable such that:

$$\begin{aligned} VAR\left(\lim_{\substack{N\to\infty \\ P_r\to 0}} z_j^{(k+1)}\right) &= \xi^{(k+1)}(x,y) \\ &= \bar{P}_r \iint_{L_k} \ell(x-u, y-v)dudv \end{aligned} \tag{3.25}$$

$$\begin{aligned} VAR\left(\lim_{\substack{N\to\infty \\ P_r\to 0}} v_j^{(k+1)}\right) &= \nu^{(k+1)}(x,y) \\ &= \bar{P}_r \iint_{L_k} \ell(x-u, y-v)\psi^{(k)}(u,v)dudv \end{aligned} \tag{3.26}$$

and their correlation coefficient is asymptotically:

$$\rho\left(\lim_{\substack{N\to\infty \\ P_r\to 0}} (z_j^{(k+1)}, v_j^{(k+1)})\right) = \sqrt{\frac{\nu^{(k+1)}(x,y)}{\xi^{(k+1)}(x,y)}} \tag{3.27}$$

Proof See Appendix B.1. ■

The following corollary is used to perform the numerical analysis for the case of non-orthogonal channels.

Corollary 3 *The average error rate (3.23) of level $k \geq 2$ is bounded as follows:*

$$\psi^{(k)}(x,y) \leq \frac{1}{2} \frac{1}{\sqrt{1+\frac{4}{N_o}\nu\left(1-\frac{\nu}{\xi}\right)}} \cdot \frac{1}{1+G\xi} \tag{3.28}$$

where $\nu = \nu^{(k)}(x,y)$ as in equation (3.26), $\xi = \xi^{(k)}(x,y)$ as in equation (3.25), and G is

$$G^{(k)}_{(x,y)} = \frac{\frac{1}{N_o}\left(1-2\frac{\nu^{(k)}_{(x,y)}}{\xi^{(k)}_{(x,y)}}\right)^2}{1+\frac{4}{N_o}\nu^{(k)}_{(x,y)}\left(1-\frac{\nu^{(k)}_{(x,y)}}{\xi^{(k)}_{(x,y)}}\right)} \tag{3.29}$$

The average error rate for $k = 1$ is given by (3.12).

Proof See Appendix B.2. ■

We now present the main result of this section.

Corollary 4 *Let α be defined as in Corollary 2. Then $\lambda \triangleq \max \psi^{(k)}(x,y)$ tends to be the fixed point of the following equation:*

$$\lambda_k \leq \frac{0.5}{\sqrt{1+0.5\alpha^2}} \cdot \sqrt{1-(1-2\lambda_{k-1})^2+\frac{1}{1+\frac{\alpha^2}{2}}} \tag{3.30}$$

For $\alpha^2/2 > 0.618$, this equation has only one positive fixed point in the range $0 \leq \lambda \leq 0.5$. This fixed point is:

$$\lambda = \frac{1+\sqrt{2+\left(1+\frac{\alpha^2}{2}\right)^{-1}}}{2\left(2+\frac{\alpha^2}{2}\right)} \tag{3.31}$$

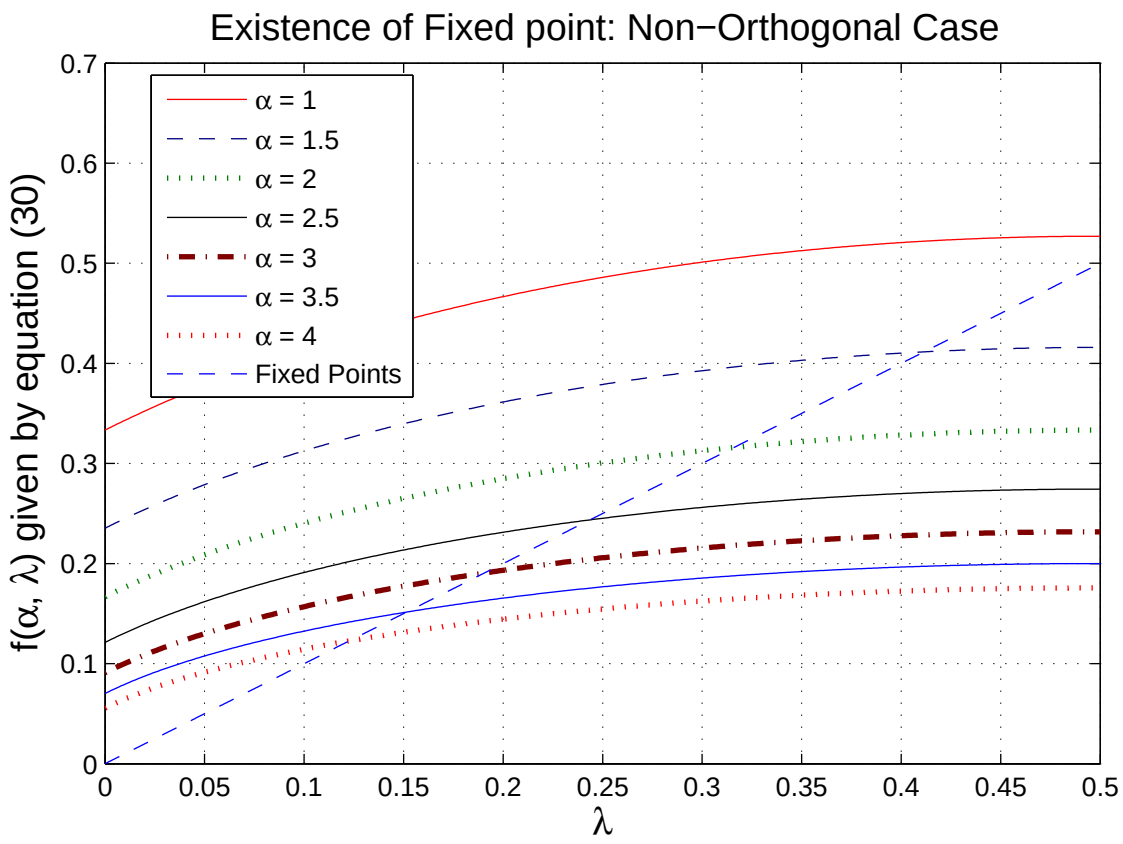


Figure 3.2: λ has a fixed point for some values of $Q(\alpha)$

For $\alpha^2/2 \leq 0.618$, the upper bound on λ is 0.5 rather than equation (3.31) (see Figure 2). For $\alpha \gg 1$, $\lambda \sim \frac{1+\sqrt{2}}{\alpha^2}$

Proof See Appendix B.3. ■

3.4 Numerical results

For both the orthogonal and non-orthogonal transmission cases we numerically solve the recursive equations that characterize the error propagation in each level.

Setting the position of the source to be the origin, the plots below show the average BER $\psi^{(k)}(x,0)$ of nodes located at coordinates $(x,0)$. For both types of transmission, the BER for level 1 is given by equation (3.12).

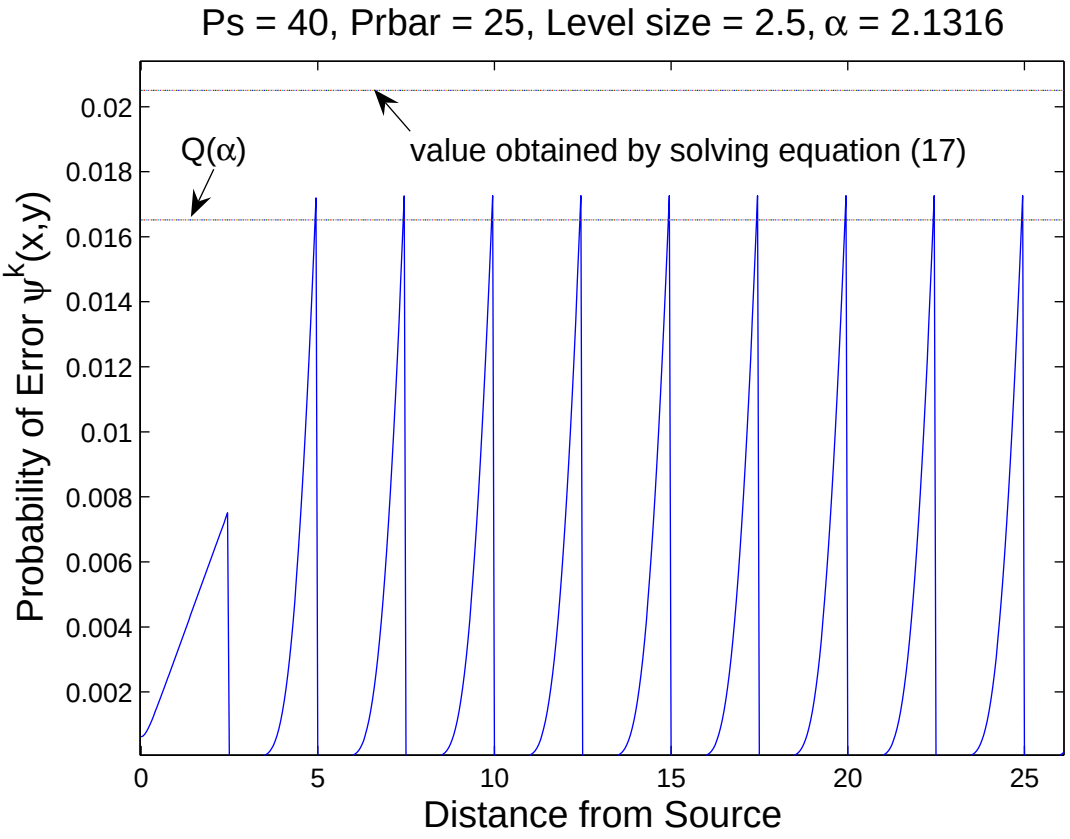


Figure 3.3: Error propagation in Orthogonal Transmission. Each tooth of the saw-tooth-like function represents the BER in a separate level.

For orthogonal transmissions the BER for subsequent levels is determined by solving equation (3.11). The numerical evaluation shows that the analytical results presented in Corollary 2 provide accurate values for the fixed point. Figure 3.3 shows that for values of $\bar{P}_r$ and level size such that $\alpha > \sqrt{\pi/2}$, the error propagation can be controlled and the worst error probability, $\psi^k(x,y)$, is close to $Q(\alpha)$. On the other hand, when $\alpha < \sqrt{\pi/2}$ error propagation is catastrophic. This is shown in Figure 3.4.

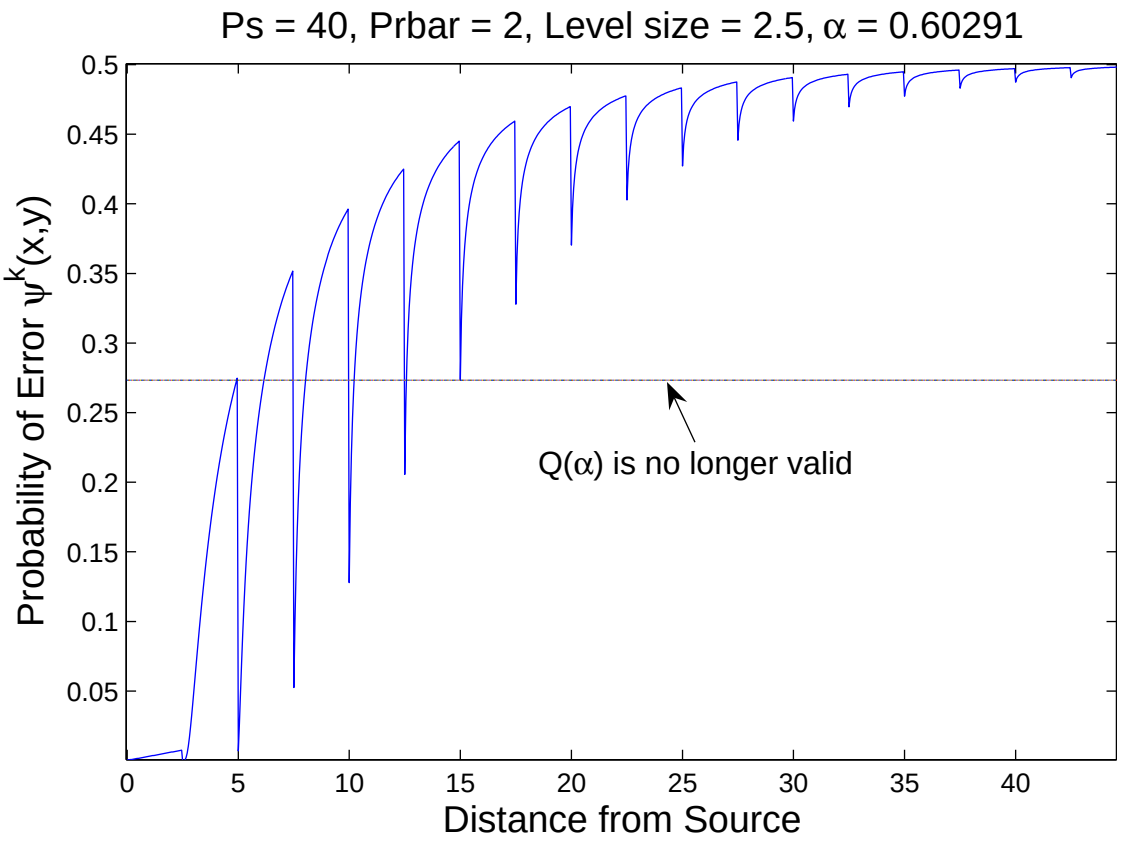


Figure 3.4: Error propagation in Orthogonal Transmission. When $\alpha < \sqrt{\pi/2}$ there is catastrophic error propagation.

In Figure 3.5 we plot the average BER as given by equation (3.28). Again, if $\bar{P}_r$ and level size are set to values such that $\alpha^2/2 > 0.618$ then the BER can be controlled. It is observed that upper bound predicted by Corollary 4 is larger than the fixed point obtained. This follows from the fact that Corollary 4 is obtained by upper bounding (3.28).

In Figures 3.6 and 3.7 we plot the BER for the orthogonal and non-orthogonal cases as given by equations (3.7) and (3.23) for different network node densities. In each case the BER is averaged over several trials to average out the effects of random Rayleigh fading and the random locations of the nodes. The figures indi-

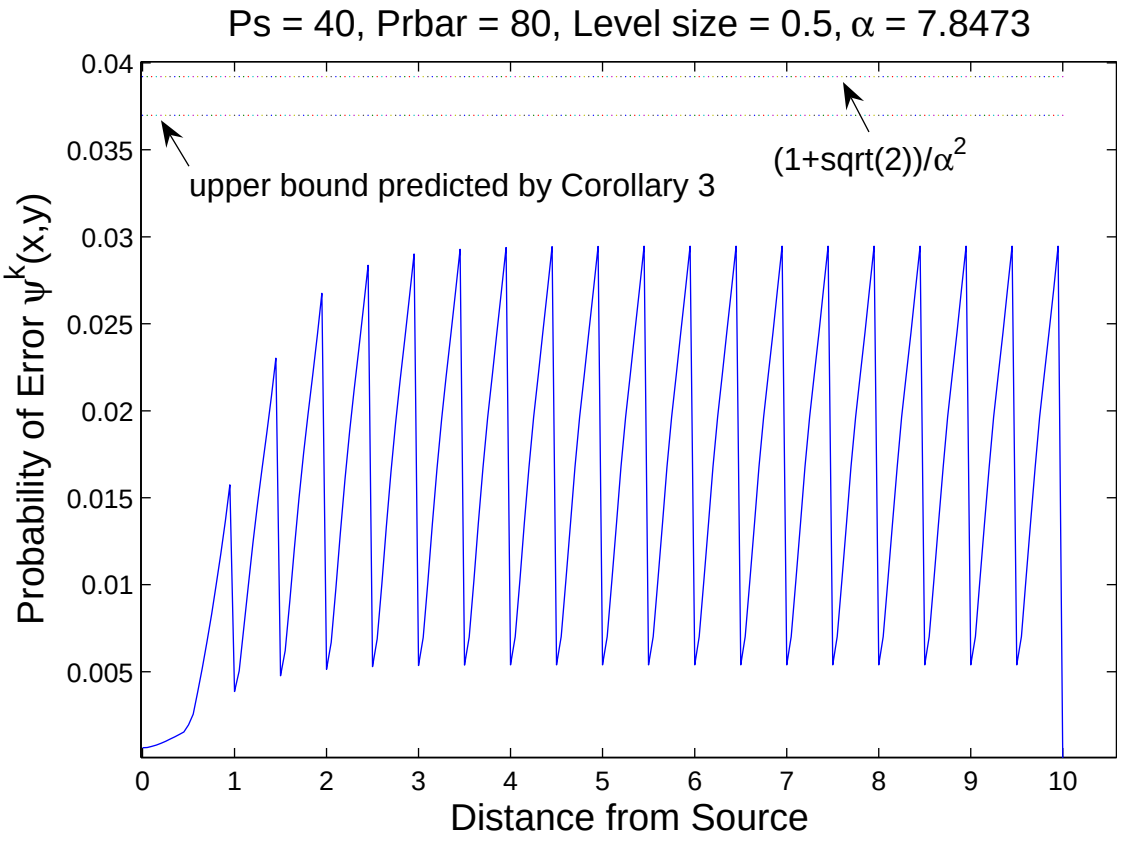


Figure 3.5: Error propagation in Non-Orthogonal Transmission. Each tooth of the saw-tooth-like function represents the BER in a separate level.

cate that the results are consistent with the BER bounds predicted by equations (3.11) and (3.28).

3.5 Summary

The asymptotic results in Corollaries 2 and 4 show that not only is the error bounded by the fixed point but that the fixed point, λ, is a nice function of the parameter α^2; where α can be interpreted as the cumulative SNR for each level. It is evident from Figure 1 that for the orthogonal transmission scheme the lower

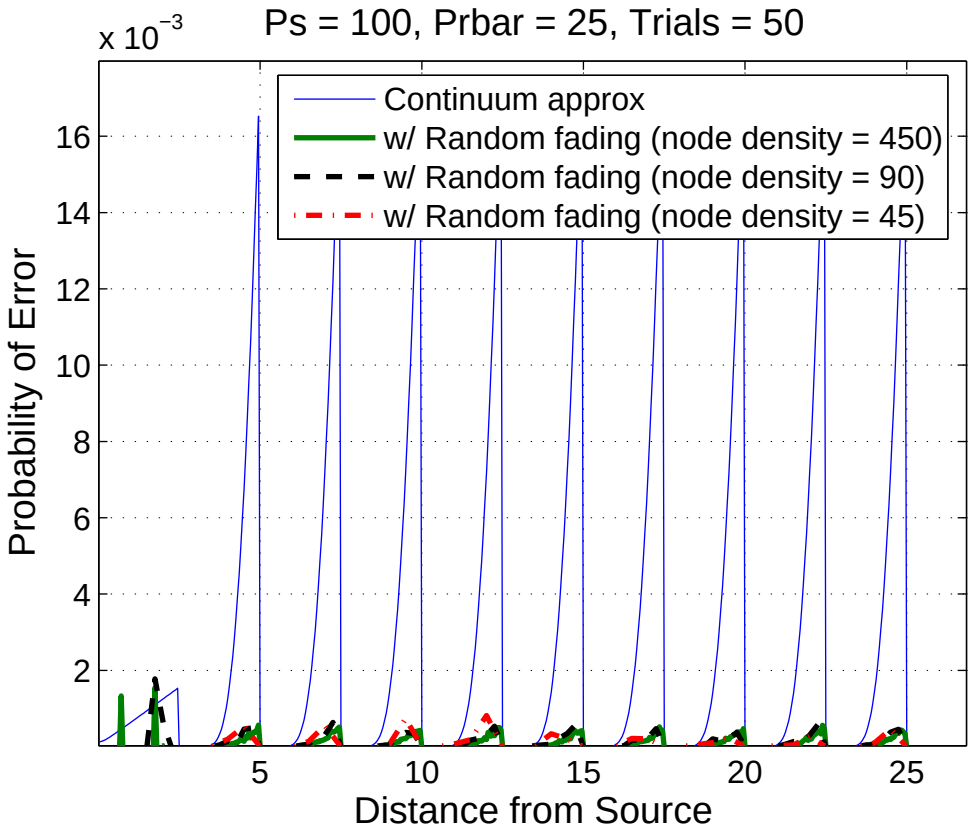


Figure 3.6: Error propagation in Orthogonal Transmission for different network densities.

bound on the BER, $Q(\alpha)$, becomes tight for $\alpha \gg 1$. This scheme requires infinite bandwidth for gaining infinite diversity over the wireless medium and it nearly matches the performance of a cooperative transmission scheme over a deterministic channel *i.e.,* an AWGN channel without fading. In contrast, the non-orthogonal transmission scheme cannot avoid fading because all cooperating nodes transmit asynchronously with respect to their carrier-phases. Here, the BER is $\rightarrow 1/\alpha^2$, which is directly proportional to $1/SNR$ as one would intuitively guess. This result should not discourage us from using the non-orthogonal transmission scheme because in terms of providing a controllable average BER performance the non-orthogonal scheme is as effective as the orthogonal scheme. Furthermore, it does

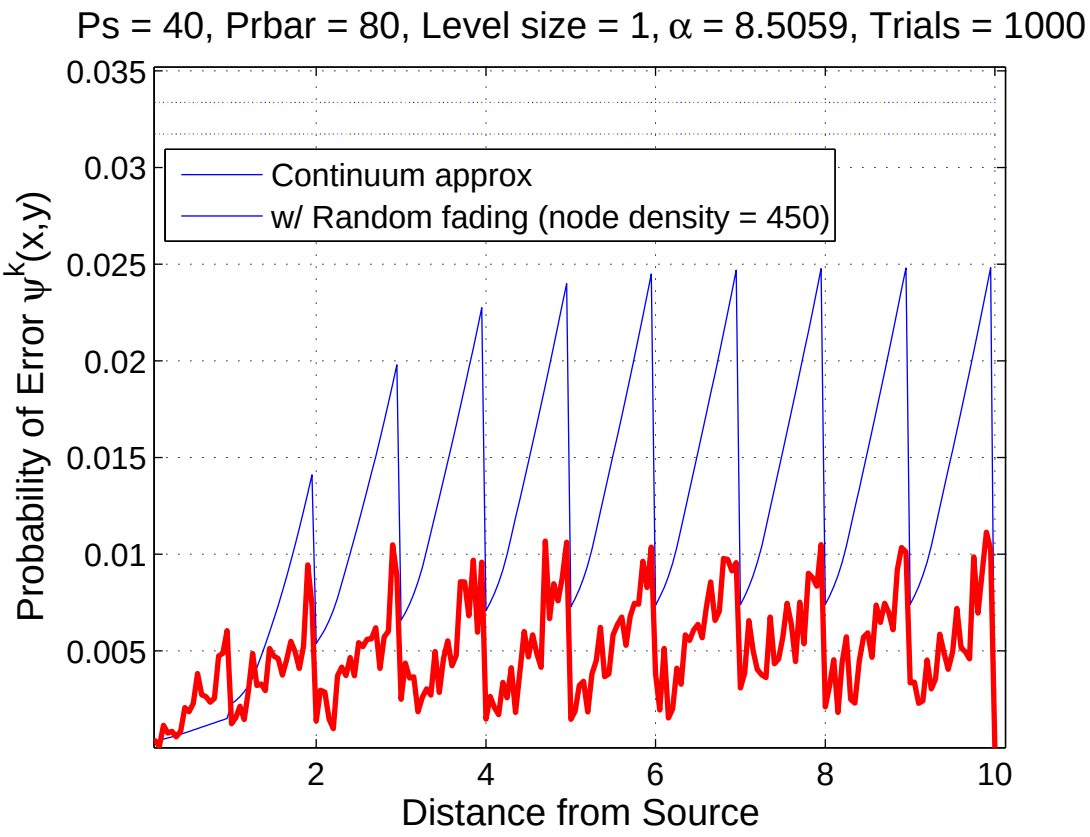


Figure 3.7: Error propagation in Non-Orthogonal Transmission for different network densities.

not require a bandwidth expansion because it allows transmissions from all cooperating nodes to effectively *collide*. The non-orthogonal scheme is also much simpler to use in a distributed network setting. To use the orthogonal scheme, the orthogonal channels have to somehow be assigned between the cooperating nodes. For the non-orthogonal scheme, on the other hand, we only require the synchronization of the nodes to the beginning of the packet they received. By using multi-carrier and/or spread spectrum transmission in each level, moderate asynchronism among the cooperating nodes can be tolerated. In this case, we expect to observe a similar scaling law for the error rates even if the transmissions of the nodes in a level are not perfectly synchronized as required by our simple model.

CHAPTER 4

CONSENSUS IN WIRELESS NETWORKS WITH INTERFERENCE

We now shift focus and consider the problem of efficiently processing network information using average consensus algorithms. Average consensus protocols have seen a recent resurgence [83, 9, 63, 70, 93, 14, 96, 55]. They can be used to compute any linear function of the network observations [94, 61] and many other applications can be built on top of them [83, 19, 60]. In wireless networks, average consensus is typically implemented using point-to-point random access scheduling. Wireless communications, however, are interference limited. What is the impact of interference due to fading on average consensus algorithms in wireless networks? It is unclear *a priori* how the schedule, either randomized or deterministic, or the transmission powers should be chosen for best performance in the face of fading. In this chapter we study the effects of interference on the averaging time and total power expenditure by considering the tradeoff between network connectivity and spatial-reuse for a circulant network. For convenience, we focus on synchronous consensus where time is slotted and a fixed number of nodes perform pairwise exchanges per slot. This still captures the essential features of asynchronous consensus with finite time updates where the number of pairwise exchanges per slot is random.

Our result shows that randomized schedules converge slowly due to interference. In contrast, well-designed schedules that mitigate interference are robust and converge faster; they also incur a lower power cost. We show the existence of an optimum transmission power for deterministic schedules which minimizes the averaging time. This operating point corresponds to the optimal trade-off between network connectivity and spatial-reuse. However, it is sufficient to use the small-

est transmission power that maintains network connectivity when minimizing the power cost.

In related work, Vanka, et al. [84] examined the effect of interference in networks on torii. Network connectivity was specified by the protocol model and communication within each disc was *broadcast* but without fading. [84] showed that for large networks, establishing long-range links increases the convergence rate for 1-torus networks. In this work we limit ourselves to pair-wise averaging protocols and do not consider broadcast variations since broadcast consensus algorithms do not converge to the initial average without acknowledgements [22]. In [85], the authors extended their results to obtain scaling laws on the convergence rate for interference-limited networks without fading.

4.1 System Model

Consider N nodes, indexed $\{0, \ldots, N-1\}$, uniformly spaced on a 1-torus (ring) of radius r. Each node is connected to K other nodes on either side of it, yielding a $2K$-regular graph. Let $\mathbf{x}(0) = [\mathbf{x}_0(0), \ldots, \mathbf{x}_{N-1}(0)]^T$ be the initial network state, where $\mathbf{x}_i(0) \in \mathbb{R}$, $i \in \{0, \ldots, N-1\}$, is a local observation known only to node i. In each iteration of the algorithm, each node exchanges its local state pairwise with each of its neighbors. These pairwise exchanges obey the following schedule.

Schedule: We consider synchronous consensus [11] where time is slotted universally. In each slot, Q node-pairs average their values in parallel according to schedule $\mathcal{S} = \{\mathcal{S}_0, \mathcal{S}_1, \ldots\}$. Let $\mathcal{S}_t$ denote the set of all (i, j) node-pairs that average pairwise in slot $t \geq 0$. We describe the schedule below (also see Fig. 4.1). We

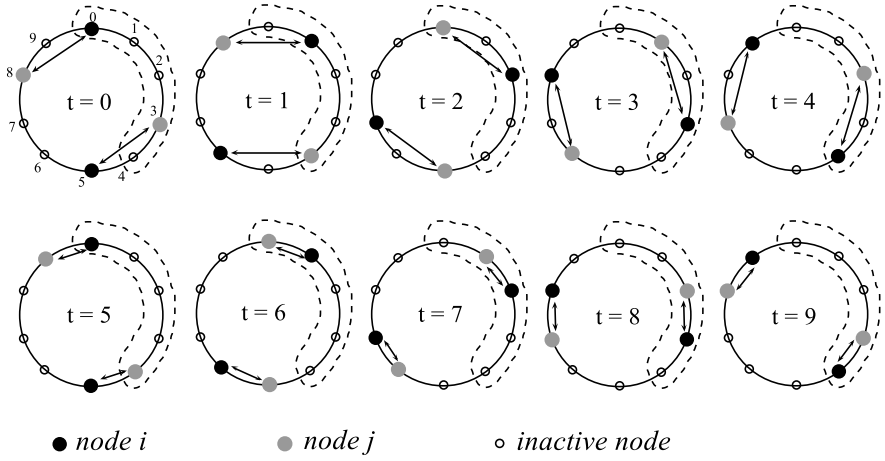


Figure 4.1: Schedule $\mathcal{S}$ for a complete round in a network with $N = 10$ nodes and $K = 2$. The nodes enclosed by the dashed line form group $\mathcal{G}_0$.

assume the following relation holds

$$N = Q(2K + 1), \tag{4.1}$$

i.e., the network can be partitioned into Q sets $\mathcal{G}_0, \ldots, \mathcal{G}_{Q-1}$, of $2K + 1 = N/Q$ nodes each, with $\mathcal{G}_0 = \{0, \ldots, \frac{N}{Q}-1\}$. Schedule $\mathcal{S}$ states that during the first $\frac{N}{Q}$ slots $t = \{0, \ldots, \frac{N}{Q}-1\}$, initiating node $i = t$ from *each* group will average pairwise with its left K^{th} neighbor (denoted j). During the next $\frac{N}{Q}$ slots $t = \{\frac{N}{Q}, \ldots, \frac{2N}{Q} - 1\}$, initiating node $i = \text{mod}(t, N/Q)$ from each group will average pairwise with its left $(K-1)^{st}$ neighbor. This process continues. Let $\mathcal{B}_t$ denote the set of the Q initiating nodes in slot t. We will see later, such a structure is preferable to a randomized schedule because it intrinsically mitigates interference. Formally,

$$\mathcal{S}_t = \{(i, j) : i \in \boldsymbol{\vartheta}^t, j = \mathcal{N}_i(\text{mod}(\lceil t/2K \rceil - 1, 2K) + 1),\ \forall i\ \}, \tag{4.2}$$

where $\boldsymbol{\vartheta}^t_q = \text{mod}(t, N/Q) + (1 + 2K)q, q = 0, \ldots, Q - 1$, $\lceil \cdot \rceil$ denotes the ceiling operator, and $\mathcal{N}_i(\ell)$ is the ℓth neighbor of node i.

In $T = \frac{NK}{Q}$ slots each node in the network can average pairwise with all its

neighbors, assuming all exchanges are successful. Q can be tuned via K by choosing a suitable transmission power. A large Q decreases T but also increases interference. The sequence $\{\mathcal{S}_0, \ldots, \mathcal{S}_{T-1}\}$ repeats every T slots and $\mathcal{S}_{mT} = \mathcal{S}_0, m \in \mathbb{Z}^+$. Finding $\mathcal{S}$ for general graphs is difficult but the idea of interference mitigation can be adapted.

Interference Model: We employ a physical model: communication from node i to j is successful in slot t if

$$\mathrm{SIR}_{(i,j)\in\mathcal{S}_t} = \frac{|h_{ij}(t)|^2}{\sum_{\ell\in\mathcal{B}_t,\ell\neq i}|h_{\ell j}(t)|^2} \geq \tau, \tag{4.3}$$

where τ is a fixed threshold. (4.3) is valid in a high SNR regime where additive noise is neglected. With Rayleigh fading, the channel gain between nodes i and j in slot t is $|h_{ij}(t)|^2 \sim \exp(1/\alpha_{ij})$ where $\alpha_{ij} = (1 + d_{ij})^{-\gamma}$, d_{ij} is the distance between the nodes and γ is the path-loss exponent. We assume that the channel coherence time is greater than the slot duration, the channels are *iid* over time slots, and channels between nodes are independent. Nodes are half-duplex and cannot simultaneously transmit and receive.

Pairwise averaging requires two-way communication. The circulant topology ensures that the interference statistics are identical in each direction of communication. Therefore, the probability of successful pairwise averaging by pair (i, j), given that they are scheduled to average in slot t, is $P(\mathrm{SIR}_{(i,j)\in\mathcal{S}_t} \geq \tau)^2$.

Consensus Updates: The network state at the end of slot t is given by $\mathbf{x}(t) = W_{\mathcal{S}_t} W_{\mathcal{S}_{t-1}} \ldots W_{\mathcal{S}_0}\mathbf{x}(0)$, where $W_{\mathcal{S}_t}$ is a random doubly stochastic matrix specified by

$$W_{\mathcal{S}_t} = I - \sum_{(i,j)\in\mathcal{S}_t} \frac{1}{2}(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T \mathbf{1}_{\{\mathrm{SIR}_{(\mathrm{i,j})},\mathrm{SIR}_{(\mathrm{j,i})}\geq\tau\}}, \tag{4.4}$$

and $\mathbf{e}_i = [0 \ldots 0\ 1\ 0 \ldots 0]^T$ is an $N \times 1$ unit vector with 1 in the i^{th} component.

Note, $W_{\mathcal{S}_t}$ is a random matrix.

4.2 Interference and Averaging Time

We characterize the averaging time for the deterministic schedule, explicitly accounting for the interference resulting from fading. Proofs are relegated to the appendix.

The ϵ-averaging time $T_{ave}(\epsilon)$, for any $0 < \epsilon < 1$, is [11]

$$\sup_{\mathbf{x}(0)} \inf_{t \geq 0} \left\{ t : P\left(\frac{||\mathbf{x}(t) - x_{ave}\mathbf{1}||}{||\mathbf{x}(0)||} \geq \epsilon \right) \leq \epsilon \right\}, \tag{4.5}$$

where $||\mathbf{x}||$ is the ℓ_2 norm of $\mathbf{x}$ and $x_{ave} = \mathbf{1}^T\mathbf{x}(0)/N$. Theorem 4, [11] bounds $T_{ave}(\epsilon)$ for general graphs:

$$\frac{0.5 \log \epsilon^{-1}}{\log \lambda_2(W^T W)^{-1}} \leq T_{ave}(\epsilon) \leq \frac{3 \log \epsilon^{-1}}{\log \lambda_2(W^T W)^{-1}}, \tag{4.6}$$

where $\lambda_2(W^T W)$ is the second largest eigenvalue of the mean averaging matrix W.

In our case, $T_{ave}(\epsilon)$ is bounded as follows. Define the random matrix $W(mT) = \prod_{i=(m-1)T}^{mT-1} W_{\mathcal{S}_i}$, for $m \in \mathbb{Z}^+$. Since the sequence $\{\mathcal{S}_0, \ldots, \mathcal{S}_{T-1}\}$ repeats every T slots, the matrices $W(mT)$ are *iid* over m. The network state at the end of every $mT = mnK/Q$ slots is $\mathbf{x}(mT) = W(mT) \ldots W(T)\mathbf{x}(0)$. Furthermore, the matrices $W(mT)$ are non-negative and doubly-stochastic. So using the result from [11], the averaging time is bounded as

$$\tilde{T}_{ave}(\epsilon) \leq \left(\frac{nK}{Q} \right) \left(\frac{3 \log \epsilon^{-1}}{\log \lambda_2(\mathbb{E}\{W^T(T)W(T)\})^{-1}} \right), \tag{4.7}$$

where λ_2 is computed in the sequel. Note that $\tilde{T}_{ave}$ is measured in multiples of the period $T = nK/Q$.

$W(T)$ is defined over $0 \leq t \leq T-1$, so it is sufficient to consider the first T slots of consensus. Let $\bar{W}_{\mathcal{S}_t} = \mathbb{E}\{W_{\mathcal{S}_t}\}$. By linearity of expectation, from (4.4),

$$\bar{W}_{\mathcal{S}_t} = I - \phi_t \sum_{(i,j)\in\mathcal{S}_t} \tfrac{1}{2}(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T, \tag{4.8}$$

where $\phi_t \triangleq P(\mathrm{SIR}_{(i,j)\in\mathcal{S}_t} \geq \tau)^2$, the probability of a successful exchange, is bounded in Theorem 2. The interference statistics experienced by node j when i transmits to it, for any pair $(i,j) \in \mathcal{S}_t$ are identical due to the choice of topology and schedule, so ϕ_t doesn't depend on specific (i,j) pairs. Moreover, during slots $N(k-1)/Q \leq t \leq Nk/Q - 1,\ k \in \{1,\ldots,K\}$, *any* active node-pair must be k hops apart (see top row of Fig. 4.1). Thus, if we decompose the first T slots in this way, then ϕ_t is solely a function of k, not t, and remains constant over each chunk of N/Q slots.

Using this, we show that $\mathbb{E}\{W^T(T)W(T)\}$ is approximately circulant. $\lambda_2(\mathbb{E}\{W^T(T)W(T)\})$ follows since the eigenvalues of circulant matrices are known. The results are summarized by the theorem below.

Theorem 2 *For schedule $\mathcal{S}$, when τ is small, the probability that node pair $(i,j) \in \mathcal{S}_t$, for $\frac{n(k-1)}{Q} \leq t \leq \frac{nk}{Q} - 1$, for $1 \leq k \leq K$, averages successfully in slot t is bounded as*

$$\sqrt{\phi_k} \geq 1 - O\left(k^{-\gamma-1}\right), \tag{4.9}$$

where γ is the path-loss exponent. The second largest eigenvalue depends on the interference parameters as

$$\lambda_2(\mathbb{E}\{W^T(T)W(T)) \lesssim 1 - 4\sin^2(\pi K/N)\left(1 - O\left(K^{-\gamma-1}\right)\right).$$

Theorem 2 specifies the dependence on the interference parameters γ and K. As K increases, network connectivity improves and by (4.1) the number of simultaneous transmissions Q decreases, reducing interference. Further, when γ is large,

severe path-loss means less interference. In both cases, $\lambda_2(\mathbb{E}\{W^T(T)W(T)\})$ decreases. The probability of successful transmission ϕ_k approaches 1 in both cases as well. The proof is provided in Appendix C.1.

4.3 Simulation Results

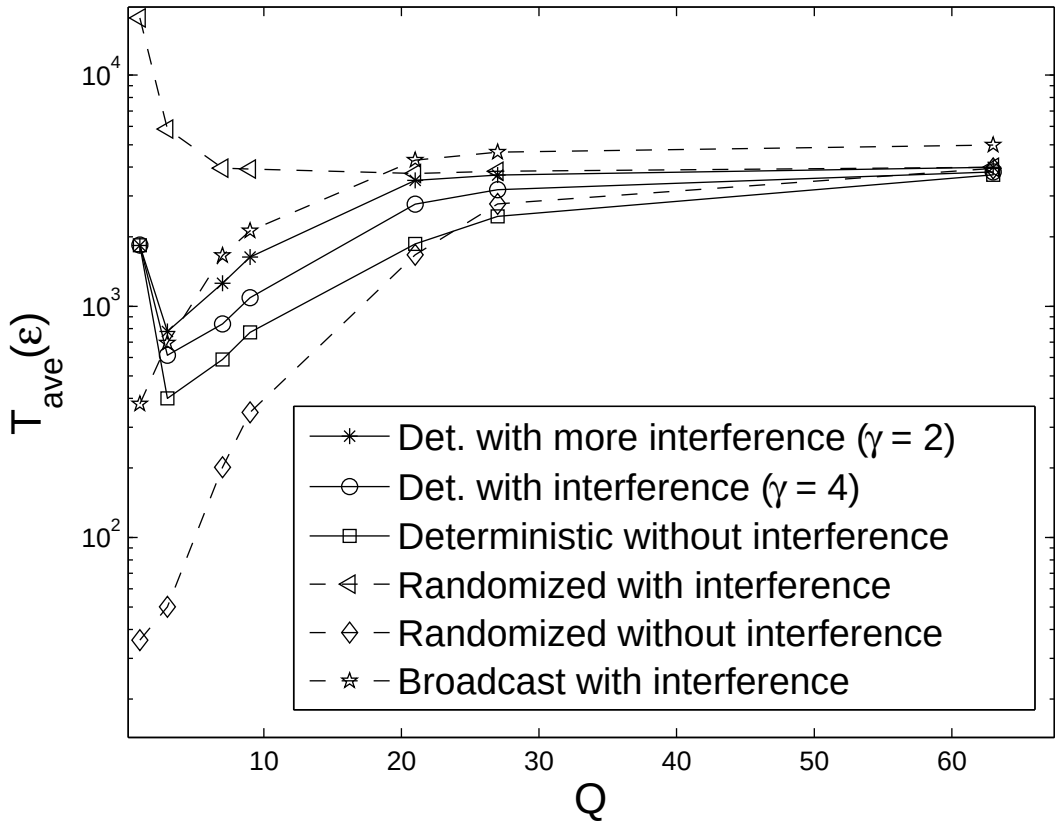


Figure 4.2: Averaging times are plotted with respect to the number of simultaneous exchanges per slot (Q) for both randomized and deterministic schedules, with and without interference.

We study the averaging time and power cost for a network with $N = 189$ nodes. The averaging cost, in terms of the total power expenditure, for the deterministic

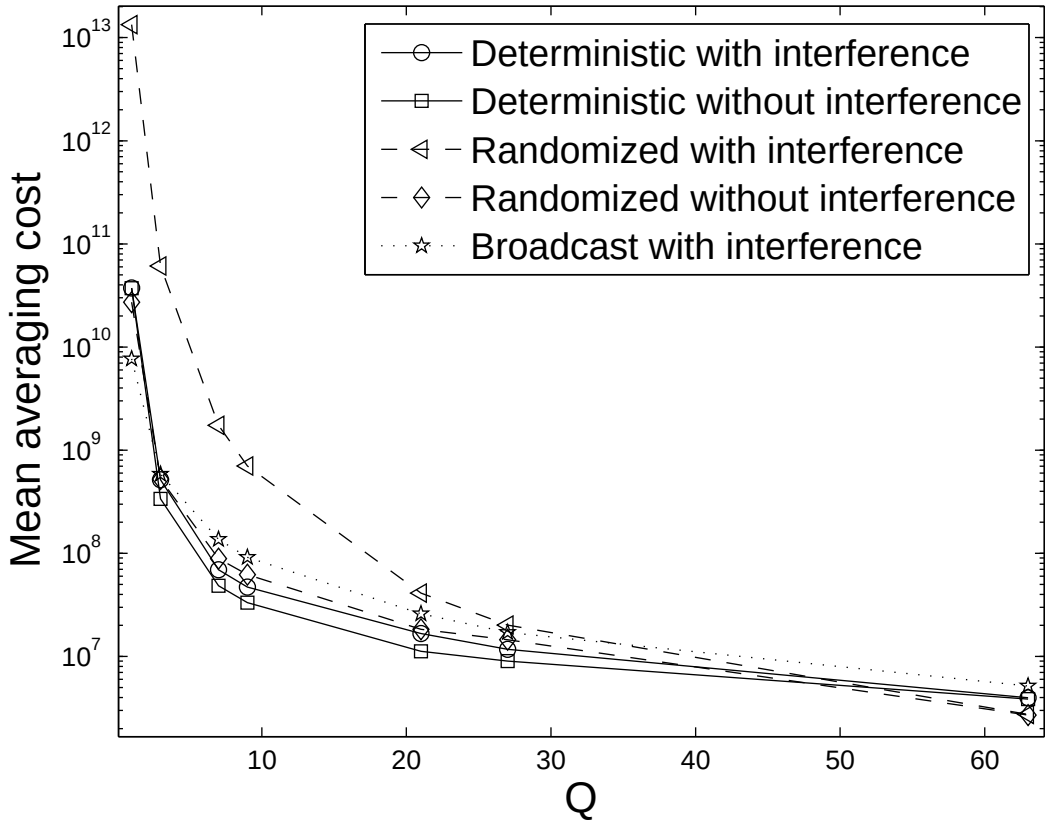


Figure 4.3: Total averaging cost in terms of power is plotted with respect to the number of simultaneous exchanges per slot (Q) for both randomized and deterministic schedules, with and without interference.

schedule is

$$C_D(d_o) = P_K \; Q \; T_{ave}(\epsilon), \tag{4.10}$$

where power $P_K \geq \tau(1 + Kd_o)^\gamma$ must hold to reach the K^{th} neighbor. The cost is parameterized by distance $d_o \triangleq d_{i,i+1}$.

For comparison, we consider a randomized schedule. It is a variation of Boyd's decentralized synchronous algorithm (§III.C.2) in [11]) where we use a $2K$-regular ring graph instead of a general network graph defined by a stochastic matrix. The

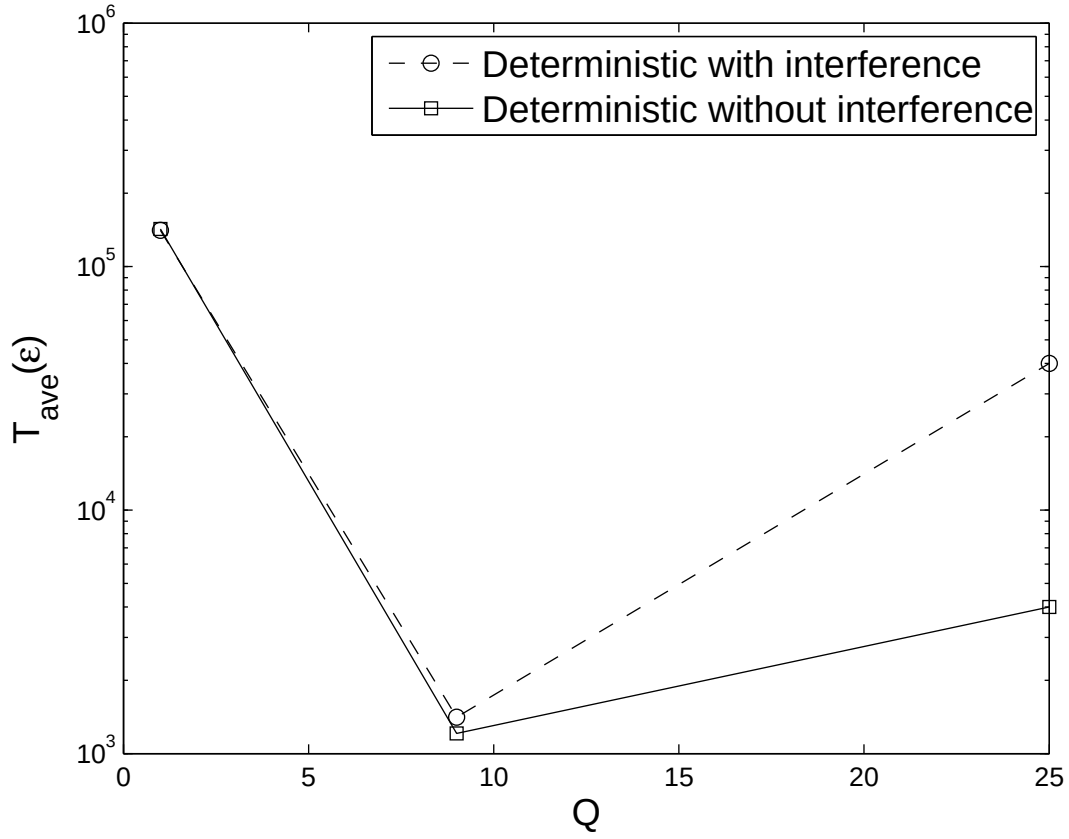


Figure 4.4: The plot shows the averaging time for a network with $N = 225$ on the 2-torus.

averaging cost for this case is

$$C_R(d_o) = P_K \; T_{ave}(\epsilon) \; \mathbb{E}(\#\text{ transmissions per slot}).$$

We plot the averaging times vs. the number of parallel transmissions Q in Fig. 4.2. Without interference, the randomized schedule exhibits a superior averaging time and $T_{ave}(\epsilon)$ scales as $O(\frac{1}{K^2})$ as expected (Theorem 8, [11]). With interference, the randomized schedule performs poorly because large transmission powers result in significant interference.

In contrast, schedule $\mathcal{S}$, with or without interference, has an optimum Q corre-

sponding to an optimal transmission power. The optimal operating point results from the interaction between network connectivity and spatial-reuse[1]: increasing K yields higher network connectivity, thus decreasing $\lambda_2(\mathbb{E}\{W^T(T)W(T)\})$; at the same time, from (4.1), this decreases the number of parallel transmissions Q and hurts spatial-reuse (see (4.7)). While fading increases the averaging time, its effect on the optimal transmission power is not large. Identical behavior is observed for a network with 225 nodes on a 2-torus (Fig. 4.4). Fig. 4.2 highlights that the interference mitigation obtained via the deterministic schedule is beneficial.

Fig. 4.3 considers the averaging cost. Schedule $\mathcal{S}$ generally displays lower cost. In all cases, the minimum cost corresponds to choosing $K = 1$, *i.e.*, the smallest transmission power is selected such that the network remains connected.

[1]This phenomenon was not captured by [84] because it assumed that whenever a node transmits, *all* its neighbors receive its state correctly. Although we do not consider broadcast channels, we plot the performance of their scheme (with fading) for comparison.

CHAPTER 5

STRUCTURED CODES FOR AVERAGE CONSENSUS

The previous chapter highlighted the impact of interference due to fading on average consensus algorithms. By considering circulant networks with a fixed number of nodes, we showed that randomized schedules converge slowly due to interference. In contrast, well-designed schedules that mitigate interference are robust and converge faster; they also incur a lower power cost. In this chapter we argue that when the available bandwidth is fixed, the expected delay in achieving a certain precision in the average value, using packet-switching communication, increases as the network size scales up. For combating these limitations in large networks we propose the use of specially structured cooperative codes. We analyze a combined source and channel coding strategy that uses a non-coherent combination of power over orthogonal sub-channels resulting in cooperative transmission for the average consensus protocol. We show that in spite of the bandwidth and power limitations, with our simple strategy the delay and precision can be kept bounded while increasing the number of participants. We support our analysis with numerical results which indicate that our strategy outperforms random access scheduling even for moderately sized networks in terms of the delay and energy required to achieve consensus.

This chapter is organized as follows. In section 5.1 we present relevant results for the average consensus algorithm. We present our model and introduce structured codes for consensus in Section 5.2. In sections 5.3 we examine the convergence and mean squared error (MSE) properties of consensus using structured codes. We present numerical results comparing the performance of structured codes for consensus versus standard wireless scheduling in Section 5.4.

5.1 Standard Average Consensus

In this section we provide a brief review of relevant prior results for the synchronous average consensus algorithm from the point of view of scalability. Consider a network of N sensors whose connectivity is defined by an underlying graph with adjacency matrix $A(t)$. The adjacency matrix is written as a function of t since the network connectivity can be time varying. The initial network state is $\mathbf{x}(0) = [\mathbf{x}_1(0), \ldots, \mathbf{x}_N(0)]^T$, where the values $\mathbf{x}_i(0)$ could represent, for example, noisy observations of some physical phenomenon. At the $(t+1)$st iteration node i receives values $\mathbf{x}_j(t)$ from all the neighboring nodes $j \in \mathcal{N}_i$, where $\mathcal{N}_i$ denotes the set of neighbors of node i, and updates its state with a linear combination of the local and incoming state values

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \sum_{j=1}^{N} a_{ij}(t)(\mathbf{x}_j(t) - \mathbf{x}_i(t)), \tag{5.1}$$

where $a_{ij}(t) \geq 0$ is the ij^{th} element of the normalized non-negative network adjacency matrix $A(t)$ with $a_{ij}(t) = 0$ if $i = j$ or if the nodes are not neighbors. The update term

$$\mathbf{u}_i(t) = \sum_{j=1}^{n} a_{ij}(t)(\mathbf{x}_j(t) - \mathbf{x}_i(t)) \tag{5.2}$$

is obtained via near-neighbor communications. In matrix notation, average consensus iterations can be written as

$$\mathbf{x}(t+1) = W(t)\mathbf{x}(t), \tag{5.3}$$

where $W(t) = (I - diag(A(t)\mathbf{1}) + A(t))$ and $\mathbf{1}$ is the all-one vector.

Consensus protocols have a natural trade-off between the number of iterations required and the accuracy achievable in consensus. This trade-off has been studied extensively for both the synchronous [62, 93] and asynchronous randomized ver-

sions of the algorithm [11, 22] and is summarized in the following result. This convenient single letter characterization of the network rate of convergence, denoted by $\lambda_2(W)$, holds if we assume that the communications between neighboring nodes are always successful and have infinite precision.

Lemma 16 *Let $\mathbb{E}W(t) = W$. Let $J = \frac{1}{N}\mathbf{1}\mathbf{1}$. Given the initial state $\mathbf{x}(0)$, the Mean Squared Error (MSE) at iteration t for consensus without quantization is*

$$MSE_{AC}(t) = \frac{1}{N}\mathbb{E}\{||\mathbf{x}(t) - J\mathbf{x}(0)||^2\} \leq \lambda_2^{2t}(W)\frac{\|\mathbf{x}(0)\|^2}{N} \tag{5.4}$$

where $\lambda_2(W)$ denotes the second largest eigenvalue of W and $\|\cdot\|$ denotes the ℓ_2 norm operator.

In addition to the idealized model that leads to Lemma 16, the communication network that supports average consensus protocols has been abstracted in several different ways in literature. [63] assumes analog communication channels where the node states are exchanged in continuous time, [93] assumes channels with additive noise, while [50] considers wired peer-to-peer networks with communication delays. The prevalent communication model for wireless networks combines asynchronous average consensus with random access protocols, *e.g.* [11, 22]. These works do not account for packet losses or rate constraints. Finite precision, and more specifically, the quantization of states was considered explicitly only recently. For example, [39] considered randomized quantized consensus algorithms under the constraint that $x_i \in \mathbb{Z}$ and bounded their convergence times; [95, 96] proposed the use of variable quantization rates and the network side information produced by the consensus algorithm to improve the delay-accuracy trade-off in rate constrained wireless transmission; and [38, 6, 13] used probabilistic quantization to achieve consensus. Finally, note that while it is possible to guarantee convergence to a

quantized consensus state, either using dithering [6, 38] or a variable encoding rate [95, 96], convergence to the *quantized average* cannot be guaranteed in the mean squared sense when the precision is finite (although the error can be made arbitrarily small by increasing the precision and communication cost).

5.1.1 Scalability

We now setup a comparison of standard average consensus with the proposed structured codes. We illustrate heuristically that standard average consensus scales poorly with the network size N because $\lim_{N\to\infty} \lambda_2(W) \to 1$. Consider the following extreme strategies for exchanging messages in the network: (1) keep the communications local or (2) try to reach distant neighbors. We examine what happens to $\lambda_2(W)$ for N large for each of these strategies.

To illustrate the (lack of) scalability of the network in the first case, consider N sensors deployed in a circle of radius r and let a node communicate with k neighbors on either side of it. We let the set of neighbors $\mathcal{N}_i,\ \forall i$ be fixed while increasing N. The adjacency matrix $A = \epsilon \cdot circ(0, \mathbf{1}_k^T, \mathbf{0}_{N-2k-1}^T, \mathbf{1}_k^T)$, where $\mathbf{1}_k$ is the all-ones vector in $\mathbb{R}^k$ and ϵ some constant, is circulant. In this case the eigenvalues of W can be computed in closed form [20]. Given that $|\mathcal{N}_i| = 2k,\ \forall i$ is kept constant while N scales up, the following lemma holds.

Lemma 17 *For the given circulant network, and k fixed, the second largest eigenvalue is*

$$\lim_{N\to\infty} \lambda_2(W) \to 1 - O\left(\frac{1}{N^2}\right). \tag{5.5}$$

Proof Using Theorem 3.2.2 from [20] and simplifying, we have $\lambda_2(W) = 1 - 2k\epsilon +$

$2\sum_{\ell=1}^{k}\epsilon\cos\left(\frac{2\pi\ell}{N}\right)$. Under $\lim_{N\to\infty}$, use $\cos x \approx 1 - x^2/2$ with $x = \frac{2\pi\ell}{N}$ to get

$$\lambda_2(W) = 1 - \frac{2\pi^2\epsilon}{N^2}\frac{k(k+1)(2k+1)}{3} \tag{5.6}$$

which scales as $1 - O(\frac{1}{N^2})$. ■

This tendency is opposite to the well known requirement that for fast convergence the Fiedler eigenvalue ($\lambda_{N-1}(L) = 1-\lambda_2(W)$) should be close to 1 (and $\lambda_2(W)$close to 0) [63]. This argument can be generalized to the case of a circulant network deployed on a torus in a straightforward manner.

To illustrate the (lack of) scalability of the protocol when nodes try to communicate to all peers, possibly far away, one can refer to the analysis of *geographic gossip* [21]. Intuitively, this approach seems to improve the algebraic connectivity and decreases the average mixing time. It was shown in [21] that in this scenario, $\lambda_2(W)$ scales as $O(1 - c/N)$ where c is some constant (the position of the nodes is irrelevant to obtain this result). Thus, also in this case, albeit with a better trend, $\lambda_2(W)$ approaches 1 as $N \to \infty$ and the convergence speed decreases. This result holds more generally for nodes deployed randomly in a unit square. The physical distance plays an even greater role if one considers the congestion related with multi-hop routing, as done in [25].

5.2 Collaborative MAC for Consensus

Motivated by the poor performance of randomized schedules in Chapter 4 and the lack of scalability to large networks as highlighted above, in this section we introduce structured codes for average consensus and discuss the assumptions that support it.

5.2.1 Assumptions

We first note that the state values $\mathbf{x}(t)$ cannot be exchanged with infinite precision because of communication constraints.

A.0. (*Quantization Model*): We consider a simple suboptimal design where all nodes use the same quantizer $\mathcal{L} = \{q_1, \ldots, q_Q\}$ with Q quantizer centroids $q_\ell \in \mathbb{R}$, in the one dimensional uniform lattice. The choice of Q is predicated on the desired precision. The range of each quantizer is $C\sigma$ where $\sigma = \max_i \sqrt{VAR[\mathbf{x}_i(0)]}$ and C is a positive constant that renders clipping errors statistically negligible. The state variables are quantized as: $\mathbf{x}_i(t) \mapsto \bar{\mathbf{x}}_i(t) \in \mathcal{L}$.

The resulting quantization error $\mathbf{v}_i(t) = \bar{\mathbf{x}}_i(t) - \mathbf{x}_i(t)$ is assumed to be uncorrelated across nodes, approximately uniform, with statistics:

$$VAR[\mathbf{v}_i(t)] = \frac{C^2}{12}\frac{\sigma^2}{Q^2}, \tag{5.7}$$

$$E\{\mathbf{v}(t)\mathbf{v}^T(t)\} = \frac{C^2}{12}\frac{\sigma^2}{Q^2}I. \tag{5.8}$$

The assumption of uncorrelated quantization error, common in considering quantization noise, is unrealistic if the quantizer is too coarse, since the quantization error in that case will neither be uniform nor uncorrelated. Hence, in all our conclusions, we assume that the quantizer has sufficient resolution for **A.0.** to be valid in practice. In any case, this fact is unrelated with the network size but rather depends on the statistics of the states.

Accordingly, the quantized consensus update is

$$\bar{\mathbf{u}}_i(t) = \sum_{j=1}^{N} a_{ij}(t)(\bar{\mathbf{x}}_j(t) - \bar{\mathbf{x}}_i(t)). \tag{5.9}$$

While node i knows the unquantized local state $\mathbf{x}_i(t)$, for analytical convenience, we require node i to use the quantized local state $\bar{\mathbf{x}}_i(t)$ during the state update.

As we will see later, this restriction does not drastically impact performance.

We make the following assumptions on the communication physical layer. Let $\tau \in \mathbb{R}$ indicate a continuous time variable.

A.1. (*Signal space*): Time is slotted in intervals of duration $T = 1$. The RF signals transmitted belong to a signal space of dimension $Q \propto BT$ complex dimensions where B is the allocated bandwidth around the carrier frequency[1]. We denote by $\{c_\ell(\tau)\}_{\ell=1}^{Q}$, the base-band complex equivalent orthonormal basis chosen to span the signal space. The signal transmitted by node i in the t^{th} iteration is

$$S_i(\tau, t) = \sum_{\ell=1}^{Q} s_i(t, \ell) c_\ell(\tau - tT), \tag{5.10}$$

where $s_i(t) = [s_i(t, 1), \ldots, s_i(t, Q)]^T$ is the vector of coordinates of the transmit signal with respect to the basis $\{c_\ell(\tau)\}_{\ell=1}^{Q}$. Details regarding the choice of coefficients $s_i(t)$ are provided later.

A.2. (*Power constraint*): Each node has a per iteration power constraint $P = \sum_{\ell=1}^{Q} |s_i(t, \ell)|^2$.

A.3. (*Incoherent channel: Fading + AWGN*): Each received signal, whose complex envelope is $R_i(\tau, t)$ is affected by an independent additive white Gaussian noise process $\mathrm{W}_i(\tau)$ with noise spectral density N_0. The channel is broadcast. Its distortion on $c_i(\tau)$ due to asynchronism and multipath, can be captured by a single independent fading coefficient, changing over the slots as block fading, denoted by $h_{ij}(t) \sim \mathcal{CN}(0, \alpha_{ij})$, where α_{ij} is the average path-loss (large scale fading). Reciprocity holds on average, *i.e.*, $\alpha_{ij} = \alpha_{ji}$.

A.4. (*Half-duplex channel*): If node i has $s_i(t, \ell) \neq 0$ for $\ell \in \{1, \ldots, Q\}$, node i cannot sense any code transmitted in the sub-space corresponding to $c_\ell(\tau)$.

[1]Dimensionality Theorem, p. 294, [92]

Based on **A.1.-A.3.**, a sufficient statistic for the received signal is $r_i(t,\ell) = N^{-1/2} < R_i(\tau,t), c_\ell(\tau - \ell T) >$. Given **A.1.-A.4.**, for $\ell \in [1,Q]$ we have

$$r_i(t,\ell) = \begin{cases} \frac{1}{\sqrt{N}} \left[\sum_{j=1}^{N} h_{ji}(t) s_j(t,\ell) + w_i(t,\ell) \right], & s_i(t,\ell) = 0 \\ 0, & \text{else.} \end{cases} \tag{5.11}$$

The received vector is $r_i(t) = [r_i(t,1), \ldots, r_i(t,Q)]^T$, $\forall i$.

5.2.2 Channel Codes

We consider channel coding and decoding strategies that do not have memory across blocks, *i.e.*, the code is a mapping

$$\bar{\mathbf{x}}_i(t) \mapsto s_i(t), \quad i = 1, \ldots, N. \tag{5.12}$$

The receiver's objective is to retrieve the update estimate $\bar{\mathbf{u}}_i(t)$ in (5.9) from the received vector $r_i(t)$, *i.e.*, the decoder is a mapping

$$r_i(t) \mapsto \bar{\mathbf{u}}_i(t), \quad i = 1, \ldots, N. \tag{5.13}$$

First consider the encoder. Based on **A.0.**, the quantized update $\bar{\mathbf{u}}_i(t) = \sum_{j=1}^{N} a_{ij}(t)(\bar{\mathbf{x}}_j(t) - \bar{\mathbf{x}}_i(t))$ can be decomposed as

$$\bar{\mathbf{u}}_i(t) = \sum_{k=1}^{Q} (q_k - \bar{\mathbf{x}}_i(t)) m_i(t,k), \tag{5.14}$$

where $q_k \in \mathcal{L}$ and $m_i(t,k)$ denotes the *network information* that node i needs to know in order to compute its consensus update:

$$m_i(t,k) = \sum_{j=1}^{N} a_{ij}(t) \delta[q_k - \bar{x}_j(t)], \tag{5.15}$$

with $\delta[x] = \mathbb{I}_{\{x=0\}}$. Essentially, $m_i(t,k)$ is a measure of the number of neighbors of node i whose quantized state is q_k. Note that of all the terms $m_i(t,k)$, $1 \leq k \leq Q$, the one that corresponds to $q_k = \bar{\mathbf{x}}_i(t)$ is always weighted by zero in (5.14).

Remark 4 *In quantized consensus, for any adjacency matrix $A(t)$, two nodes whose states fall in the same quantization bin at iteration t do not need to communicate during that iteration. In practice this is enforced by the half-duplex constraint.*

Furthermore, we will now show that the network information $m_i(t,k)$ can be embedded in a code that delivers it at each node directly and collaboratively. The multiple access scheme and the codes for $m_i(t,k)$ are as follows. We choose the coefficients $s_i(t,\ell)$ as

$$\mathbf{x}_i(t) \mapsto \bar{\mathbf{x}}_i(t) \mapsto s_i(t,\ell) = e^{j\phi_\ell}\delta[q_\ell - \bar{\mathbf{x}}_i(t)],\ \ell = 1\ldots,Q, \tag{5.16}$$

where $q_\ell \in \mathcal{L}$ and $\phi_\ell \sim \mathcal{U}[0,2\pi)$.

The codes can be described as follows. Predicated on the desired precision, each consensus iteration allows the transmission of Q orthogonal waveforms. All nodes employ the coding scheme given in (5.16) which may result in cooperative transmissions: if a node's quantized state satisfies $\bar{\mathbf{x}}_i(t) = q_\ell$, then it should transmit the ℓ^{th} waveform. An illustration of the transmission scheme is shown in Fig. 1.

Lemma 18 *With these codes, when the adjacency matrix coefficients are equal to $a_{ij} = \alpha_{ij}/N$, the Maximum Likelihood (ML) estimate of $m_i(t,\ell),\ \ell = 1,\ldots,Q$ in (5.15) is:*

$$\hat{m}_i(t,\ell) = |r_i(t,\ell)|^2 - \frac{N_0}{N}. \tag{5.17}$$

Proof Let $a_{ij} = \frac{\alpha_{ij}}{N}$ for $i \neq j$. The quantity we want to estimate given $r_i(t,\ell)$ is

$$m_i(t,\ell) = \frac{1}{N}\sum_{j=1}^{N} \alpha_{ij}\delta[q_\ell - \bar{\mathbf{x}}_j(t)]. \tag{5.18}$$

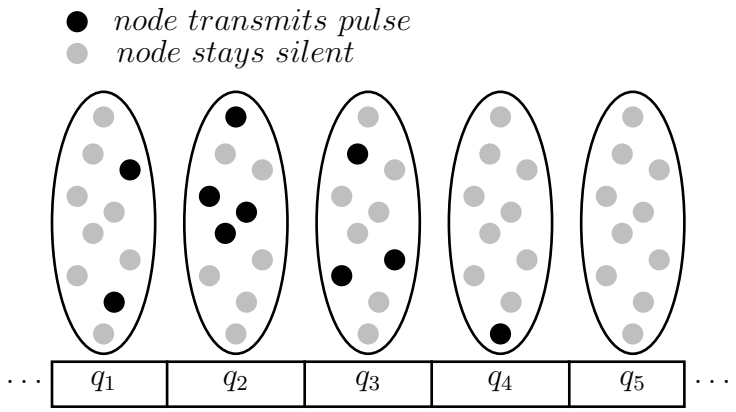


Figure 5.1: Multiple access code: A network with 10 nodes is deployed randomly in an elliptical field. Let $Q = 5$. The code specifies that each node $i = 1, \ldots, 10$, will transmit waveform ℓ corresponding to $q_\ell = \bar{\mathbf{x}}_i(t)$. The figure is shown for an arbitrary realization of $\bar{\mathbf{x}}(t)$.

Given $s_j(t, \ell)$, from (5.11) we have that

$$r_i(t, \ell) \sim \mathcal{CN}\left(0, \frac{1}{N}\sum_{j=1}^{N} \alpha_{ij}|s_j(t, \ell)|^2 + \frac{N_0}{N}\right). \tag{5.19}$$

From the coding scheme (5.16), $|s_j(t, \ell)|^2 = \delta[q_\ell - \bar{\mathbf{x}}_j(t)]$. Substituting this in (5.19) we get that $r_i(t, \ell) \sim \mathcal{CN}(0, m_i(t, \ell) + \frac{N_0}{n})$. Then the ML estimate is given by:

$$\hat{m}_i(t, \ell) \;\; = \;\; \arg\max \frac{\exp\left\{-|r_i(t, \ell)|^2/(m_i(t, \ell) + \frac{N_0}{n})\right\}}{\pi(m_i(t, \ell) + \frac{N_0}{n})}.$$

The result follows upon simplification. ■

The encoding strategy is given in (5.16). Now consider the decoder. To decode, each node i constructs

$$\hat{\mathbf{u}}_i(t) = \sum_{\ell=1}^{Q} (q_\ell - \bar{\mathbf{x}}_i(t))\hat{m}_i(t, \ell) \tag{5.20}$$

from the received signal vector $r_i(t)$. Based on Lemma 18, $\hat{\mathbf{u}}_i(t)$ is the estimate of $\bar{\mathbf{u}}_i(t)$. Now the usual consensus update of the form (5.1) can be applied using quantized states.

We note that in the aforementioned scheme, the ML estimate of $m_i(t,\ell)$ is obtained by using only one received signal sample. An extremely simple way of reducing the error in estimating the network information is to use repetition coding. Suppose a $(\Psi, 1)$ repetition code is employed by all nodes. Then the code corresponding to the coefficient $s_i(t,\ell)$ is transmitted by each node Ψ times. The cost of using the repetition code is a bandwidth expansion on the order of Ψ per iteration. Repetition coding is summarized by the following corollary. The proof is trivial and is omitted.

Corollary 5 *When a $(\Psi, 1)$ repetition code is used, the ML estimate of $m_i(t,\ell)$ is*

$$\hat{m}_i(t,\ell) = \frac{1}{\Psi}\sum_{s=1}^{\Psi} |r_{is}(t,\ell)|^2 - \frac{N_0}{n} \tag{5.21}$$

where $r_{is}(t,\ell)$ denotes the s^{th} sample of $r_i(t,\ell)$.

In the remainder of this chapter we analyze the performance of these data driven codes and verify that it scales with the network size.

5.3 Performance Analysis

First we show that the data driven consensus architecture is unbiased with respect to quantized consensus and it achieves low MSE. Next, since the average consensus algorithm does not converge to the true mean x_{ave} in general, we study the MSE achieved at an iteration t.

5.3.1 Convergence in Expectation

We show that in terms of the average behavior, our algorithm behaves exactly like the standard average consensus algorithm using *quantized* states and therefore converges in expectation.

Lemma 19 *For any given initial state* $\mathbf{x}(0)$*:*

$$E\{\hat{m}_i(t,\ell)\} = m_i(t,\ell),\ \forall \ell.$$

Therefore, the multiple access coding method proposed on average tends to the same result as quantized consensus.

Proof $r_i(t,\ell)$ is circularly symmetric complex gaussian (5.19). So $|r_i(t,\ell)|^2$ is exponential with parameter $(\frac{1}{N}\sum_{j=1}^{N}\alpha_{ij}\delta[q_\ell - \bar{\mathbf{x}}_j(t)] + N_0/N)^{-1}$. Noting $a_{ij} = \alpha_{ij}/N$, irrespective of the repetition code order Ψ, it follows that

$$E\{\hat{m}_i(t,\ell)\} = \sum_{j=1}^{N} a_{ij}\delta[q_\ell - \bar{\mathbf{x}}_j(t)] = m_i(t,\ell) \tag{5.22}$$

as desired. ■

5.3.2 Mean Square Error

From our construction it is clear that one iteration entails a delay equal to the interval T that is the duration of our orthogonal basis $\{c_\ell(\tau)\}_{\ell=1}^{Q}$. Since the data driven consensus algorithm accumulates quantization errors and channel errors, the mean squared error (MSE) after a certain delay is an essential performance metric.

For our scheme, the MSE can be decomposed into contributions from three separate terms: (1) the convergence error $\mathbf{x}(t) - J\mathbf{x}(0)$, (2) the quantization error $\mathbf{x}(t) - \bar{\mathbf{x}}(t)$, and (3) the channel error $\mathbf{e}(t) = \hat{\mathbf{u}}(t) - \mathbb{E}\{\hat{\mathbf{u}}(t)\}$. Note that quantization error (by **A.0.**) and channel error (by definition) are zero mean.

We examine the asymptotic behavior of the MSE for a circulant topology. Consider N nodes deployed in a circle such that $a_{ij} = \alpha_{ij}/N$, as specified in Lemma 18, with $\alpha_{ij} = \mathcal{K}(d^* + d_{ij})^{-\gamma}$. γ is the path loss exponent, d_{ij} is the distance between the nodes, and d^* and $\mathcal{K}$, related by $\mathcal{K} = (\frac{1}{d^*})^{-\gamma}$, are modeling parameters that take into account the carrier frequency, the scattering environment and antennae gains. Due to symmetry in path loss, $A = \frac{1}{N} circ(0, \alpha_{1,2}, \ldots, \alpha_{1,(N+1)/2}, \alpha_{1,(N+1)/2}, \ldots, \alpha_{1,2})$ for N odd (the case for N even is similar). First we obtain an expression for $\lambda_2(W)$ for the given network topology.

Lemma 20 *For the given topology,* $\lim_{N\to\infty} \lambda_2(W) \to 1 - \frac{c\pi^2}{6} < 1$.

Proof From [20], using $N \to \infty$, we get-

$$\lambda_2(W) = 1 - \frac{4\pi^2}{N^3} \sum_{r=1}^{k'} r^2 \alpha_{1,r} \leq 1 - \frac{\pi^2(N^2-1)}{6N^2} \min_r \alpha_{1,r} \tag{5.23}$$

$$= 1 - \frac{c\pi^2}{6} \tag{5.24}$$

where $k' = (N-1)/2$ and $c = \min_r \alpha_{1,r} = \mathcal{K}(d^* + d_{1,\frac{n-1}{2}})^{-\gamma}$. ■

Note that because of our selection of entries a_{ij}, the behavior of $\lambda_2(W)$ is different from that given in Lemma 17 for a similar topology. Then the asymptotic average MSE characterizing the tradeoff between the allocated communication resources and precision in consensus is given by the following lemma.

Lemma 21 *For the given circulant network, as $N \to \infty$, the average MSE at a*

finite iteration t is

$$\begin{aligned}\lim_{N\to\infty} MSE(t) &\leq \lambda_2^{2t}(W) \lim_{N\to\infty} \frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} + O(f(\bar{\Delta})) \\ &\quad + \frac{2\xi}{\Psi} \sum_{i=t}^{1} \lambda_2^{2(i-1)}(W) MSE(t-i) \\ &= \frac{O(f(\bar{\Delta}))(1-\lambda_2^2(W))}{1-\lambda_2^2(W)-\frac{2\xi}{\Psi}} \\ &\quad + \left(\lim_{N\to\infty} \frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} - \frac{\frac{2\xi}{\Psi} O(f(\bar{\Delta}))}{1-\lambda_2^2(W)-\frac{2\xi}{\Psi}} \right) \left(\lambda_2^2(W) + \frac{2\xi}{\Psi} \right)^t\end{aligned}$$

where $\lambda_2(W) \leq 1 - \frac{c\pi^2}{6} < 1$ *with c some positive constant,* $f(\bar{\Delta}) = \Delta^2/12$ *where* Δ *is the quantization interval of the quantizer,* Ψ *is the order of the repetition code, and constant* ξ *satisfies the conditions*

1. $\xi < \frac{\Psi}{2}(1-\lambda_2^2(W))$
2. $\xi \leq \frac{\Psi}{2} \left(\lim_{N\to\infty} \frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} (1-\lambda_2^2(W) \right) / \left(\lim_{N\to\infty} \frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} + O(f(\bar{\Delta})) \right)$.

Proof See Appendix D.1. ■

We note that by allocating greater precision, $f(\bar{\Delta})$ can be made arbitrarily small, thus making the quantization error term negligible. Furthermore, the recursive error term can also be made arbitrarily small by increasing the order of the repetition code Ψ. Increasing the precision or the order of the repetition code essentially amounts to an expansion in the required bandwidth. Hence, as we allocate greater communication resources, *i.e.*, bandwidth, the performance of the data driven codes for an infinitely large network is captured by the following corollary.

Corollary 6 *For the given circulant network, in the case of bandwidth expansion,*

the average MSE at a finite iteration t is

$$MSE(t) \leq \lambda_2^{2t}(W) \lim_{N\to\infty} \frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}. \tag{5.25}$$

Interestingly, this is equivalent to the average MSE of standard average consensus (c.f. Lemma 16).

On the other hand, if we operate in the regime of finite bandwidth, we incur a penalty as indicated by Lemma 21. This analysis allows us to conclude the main result of this section, *i.e.*, that it is possible to achieve a bounded average MSE irrespective of the number of nodes. Therefore, performing average consensus in dense networks with finite precision is possible in spite of limited bandwidth.

Remark 5 *Using ideas from the proof of the asymptotic MSE in Lemma 21, we can show convergence of the data driven algorithm for a finite network ($N < \infty$). Let us define a measure of the overall average deviation of the quantized states from the mean at a given iteration by $\mathbb{E}\{||\boldsymbol{\beta}(t)||^2\}/N$ where $\boldsymbol{\beta}(t) \triangleq \bar{\mathbf{x}}(t) - J\mathbf{x}(t)$ (note, this is not the same as the MSE). Then under similar conditions as Lemma 21 we can show that the average deviation converges to a floor.*

Corollary 7 *For $N < \infty$, the average deviation tends to a floor* $\frac{1}{N}\mathbb{E}\{||\boldsymbol{\beta}(t)||^2\}$

$$\leq \frac{b(1-\lambda_2^2(W))}{1-\lambda_2^2(W)-\frac{2\xi'}{\Psi}} + \left(\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} - \frac{\frac{2\xi'}{\Psi}b}{1-\lambda_2^2(W)-\frac{2\xi'}{\Psi}}\right)\left(\lambda_2^2(W)+\frac{2\xi'}{\Psi}\right)^t \tag{5.26}$$

where $b = O\left(\frac{1}{N}\right) + O\left(\frac{N_0}{\Psi N}\right) + O\left(\frac{N_0^2}{\Psi N^2}\right) + f(\bar{\Delta})\left(1+\sum_{l=t}^{1}\lambda_2^{2\ell}(W)\right)$ and constant ξ' satisfies the conditions

1. $\xi' < \frac{\Psi}{2}(1-\lambda_2^2(W))$
2. $\xi' \leq \frac{\Psi}{2}\left(\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}(1-\lambda_2^2(W)\right) / \left(\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}+b\right)$.

Proof *See Appendix D.2.* ■

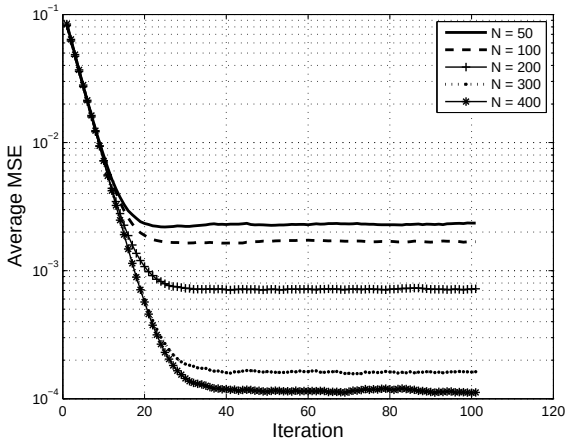
Remark 6 *Because of the difficulty of computing the constants ξ and ξ' exactly, the upper bounds on the asymptotic MSE and the average deviation of the quantized states cannot be obtained explicitly. However, in the next section we validate that the behavior of these quantities numerically.*

5.4 Numerical Results

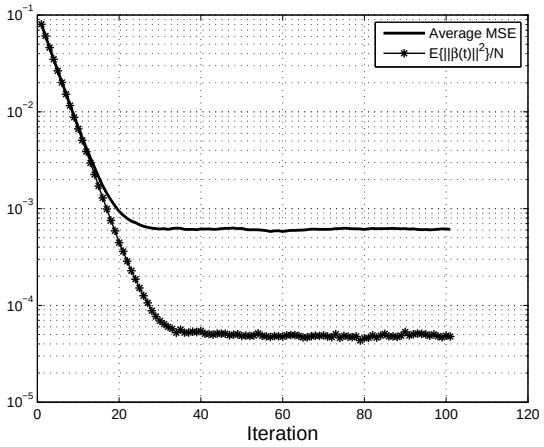
In this section we verify the scalability of data driven scheduling and compare its performance to that of standard average consensus which is implemented using packet-switched wireless communications.

The invariance of the performance of data driven scheduling with respect to network size is by now perhaps intuitive. This is shown in Fig. 5.2(a) for the circulant network topology described in Section 5.3 using $Q = 128$. The figure also shows an interesting phenomenon - as the network size increases, the achievable average MSE decreases. Fig. 5.2(b) numerically verifies that the average deviation of the quantized states tends to a floor for a finite network.

We will now show the main result of this section - if we wish to achieve a certain average MSE in consensus, then the proposed data driven scheduling can achieve it in less time than standard average consensus. In order to make a fair comparison, we require knowledge of the optimal scheduling policy for exchanging packets in the network when standard consensus is used. We refer to optimality in the sense of the fastest *average* time to convergence. Accordingly, we employ



(a)



(b)

Figure 5.2: (a) The rate of convergence of the algorithm using the data driven codes is invariant with respect to the network size N. (b) The average deviation of the quantized states tends to a floor for $N = 100$ and $Q = 128$.

the deterministic scheduling for standard consensus from Chapter 4, using the parameters that correspond to the optimal operating point determined numerically in Fig. 4.2.

The channel model is assumed to be identical for both scheduling strategies and is given by assumptions **A**.**3**. and **A**.**4**.. Briefly, the channels are half-duplex

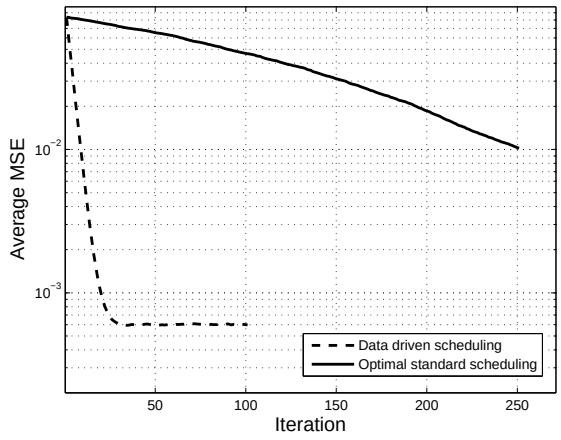


Figure 5.3: *Average MSE comparison*: MSE vs. iteration is plotted for $N = 189$ and 100 consensus iterations showing that consensus iterations using data driven scheduling are completed in significantly less time.

and the channel between nodes i and j is represented by a single coefficient $h_{ij} \sim \mathcal{CN}(0, \alpha_{ij})$ where $\alpha_{ij} = \mathcal{K}(d^* + d_{ij})^{-\gamma}$, as in Section 5.3. Data driven scheduling is simulated with the following parameters. 189 nodes are located uniformly on a circle. Transmission powers are unity and the transmit SNR at each node is $1/N_0 = 30dB$. Path loss exponent $\gamma = 4$ and a quantizer with $Q = 128$ bins is used to quantize the states. A $\Psi = 1$ order repetition code is used. Monte Carlo trials are used to compute the average MSE. Standard average consensus is simulated using the optimal deterministic schedule. However, states are allowed to be exchanged with infinite precision because it is well known that quantized consensus does not converge in the mean square sense [11]. Therefore, the actual performance of standard scheduling will be worse than shown here. Recall from Section 5.3 that each iteration of data driven scheduling is completed in time interval $T = 1$ where T is the duration of the orthogonal basis $\{c_\ell(\tau)\}_{\ell=1}^{Q}$. It is implicit that each time slot is of duration T. This comparison is shown in Fig. 5.3: data driven scheduling converges much faster than the packet-switched schedule.

CHAPTER 6

CONCLUSIONS

This thesis shows that cooperation at the physical layer can significantly improve performance with relatively small coordination overheads by exploiting the broadcast nature of the wireless channel. In the context of communication architectures for large scale wireless networks which are prone to network congestion and where routing is prohibitively costly, we examined physical layer protocols for fast intra-network broadcast communication and structured cooperative codes for fast in-network data processing through average consensus algorithms.

In the first part of the thesis we considered physical layer decode and forward cooperative protocols for broadcast. Protocols utilizing several techniques like power control, SIC, use of side-information, and interference alignment were proposed. These protocols were shown to be capable of broadcasting the content of multiple users simultaneously in the network in an efficient manner, especially in fading environments. Through detailed analysis of the two user case we determined necessary conditions for the successful broadcast in the general case. The performance improvement in terms of latency over non-cooperative packet-switched flooding was verified numerically. Finally, in the case of single user broadcast with decoding errors, it was shown that by judicious selection of network parameters, error propagation in the information flow can be controlled.

We argued that standard packet-based average consensus gossiping does not scale in dense wireless networks due to network congestion. In the second part of the thesis we proposed a solution to this problem by considering structured codes which utilize cooperative transmission through data driven scheduling. We analyzed its performance over a non-coherent AGWN-MAC channel. We showed that

Table 6.1: Selected publications related to the dissertation.

Chapter	**Selected Submissions/Publications**
2-3	Shrut Kirti, Anna Scaglione, and Bhaskar Krishnamachari, *Cooperative Broadcast in Dense Wireless Networks,* submitted to IEEE Transactions on Networking (under revision). Anna Scaglione, Shrut Kirti, and Birsen Sirkeci-Mergen, *Error Propagation in Dense Networks with Cooperation,* Proc. Intl. Symposium on Information Processing in Sensor Networks (IPSN '06), Nashville, TN, April 2006.
4-5	Shrut Kirti and Anna Scaglione, *Consensus in Wireless Networks with Interference,* submitted to IEEE Communication Letters. Shrut Kirti, Anna Scaglione, and Robert J. Thomas, *A Scalable Wireless Communication Architecture for Average Consensus,* Proc. 46th IEEE Conference on Decision and Control (CDC'07), New Orleans, LA, USA, Dec 12-14, 2007.

the algorithm converges in the mean and we characterized the MSE performance. We showed that adding infinite values with finite precision is possible in spite of the bandwidth limitations. Through numerical simulations we further showed that our codes outperform standard average consensus in terms of the time required to achieve a certain MSE in the consensus value.

Selected publications corresponding to the chapters mentioned above are listed in Table 6.1.

6.1 Future Directions

There exists a vast body of literature now that studies the capacity of multi-hop wireless networks with multiple source-destination pairs [25, 64]. None of these works, however, consider the latency involved in communication; while packet delivery is assured, no delay guarantee is provided. Several other works [8, 58, 23, 76, 1] do consider the delay-throughput tradeoff for such networks with multiple unicast flows. On the other hand, works like [42, 71, 43, 34], study the latency in broadcasting only a single source. Therefore, one avenue for further research is the delay-throughput analysis of the cooperative protocols for the simultaneous broadcast of multiple sources proposed in this thesis. Further research directions include the issue of large scale network synchronization, which we assumed exists, and a thorough analytical characterization of the opportunistic interference alignment (OIA) protocol. As regards the physical layer architecture for average consensus, it will be interesting to see it implemented and tested in actual hardware.

APPENDIX A

PROOFS FOR CHAPTER 2

A.1 Proof of Lemma 2

The proof follows the same geometric arguments used in the proof of Lemma 1 so we only highlight the differences. Let the current network state be $\boldsymbol{\mathcal{L}}_k^{(1)}$.

(a) If $x \in \boldsymbol{\mathcal{T}}_k^i,\ i = 1, 2, b$, *i.e.*, decoded some symbol in slot k, then x cannot transmit in slot $k+1$ due to the half-duplex constraint (A.3) and $x \notin \boldsymbol{\mathcal{T}}_{k+1}^i,\ i = 1, 2, b$. Hence, the nodes that decode anything in slot $k+1$ must lie in: $x < r_{k-1}$, $r_k < x < L - r_k$, or $x > L - r_{k-1}$. We show that no information flow is possible for $x > L - r_{k-1}$ in slot $k+1$. By symmetry, the result is also applicable to $x < r_{k-1}$, proving part (a).

Consider a relay at any coordinate $x > L - r_{k-1}$. Any such relay must have already forwarded s_2 in a prior slot (see Remark 2). By assumption A.4, x is therefore only interested in decoding and forwarding s_1. However, since $\boldsymbol{\mathcal{T}}_k^2 = [L - r_k, L - r_{k-1})$ is physically closer to x than $\boldsymbol{\mathcal{T}}_k^1$, symbol s_2 will be received with greater power than s_1. Its SIC receiver will therefore, decode s_1 second. We will show that the received SNR is insufficient to decode s_1 for our choice of x. Now,

$$\begin{aligned} &\mathrm{SNR}_{k+1}^1(x) \\ &= \frac{P_r \rho}{N_o} \int_{r_{k-1}}^{r_k} \ell(u, x) du < \frac{P_r \rho}{N_o} \frac{\Delta_k}{2} \left(\ell(r_k, x) + \ell(r_{k-1}, x) \right) \end{aligned}$$

where $\Delta_k = r_k - r_{k-1}$ and the upper bound is obtained by replacing the integral with the Riemann sum. Set $x = L - r_{k-1} + \epsilon$ for any $\epsilon > 0$. Since path loss is monotonic, this value maximizes $\mathrm{SNR}_{k+1}^1(x)$. By substituting this value, using the

definition of the path loss function, and noting that $r_{k-1} \leq r_k \leq \frac{L}{2}$, we can further upper bound the SNR

$$\mathrm{SNR}_{k+1}^1(x) < (P_r\rho/N_o)\Delta_k(L - 2r_{k-1} - \Delta_k + d_0 + \epsilon)^{-2}.$$

An upper bound on SNR is sufficient for showing $\mathrm{SNR}_{k+1}^1(x) < \tau$. Rewrite this as

$$(L - 2r_{k-1} - \Delta_k + d_0 + \epsilon)^2/\Delta_k > P_r\rho/\tau N_o. \tag{A.1}$$

To prove this, we will upper bound Δ_k. Since $\boldsymbol{\mathcal{L}}_k = \boldsymbol{\mathcal{L}}_k^{(1)}$, from part (b) we know that

$$\begin{aligned} r_k &< \frac{r_{k-1} + r_{k-2}}{2} + \frac{1}{2}\sqrt{\Delta_{k-1}^2 + \frac{4}{\frac{\tau N_0}{P_r\rho\Delta_{k-1}} + \psi(x')}} - d_0 \\ &< \frac{r_{k-1} + r_{k-2}}{2} + \frac{1}{2}\sqrt{\Delta_{k-1}^2 + 4P_r\rho\Delta_{k-1}/\tau N_0} - d_0 \end{aligned}$$

since $\psi(x') > 0$. Rearranging terms,

$$\begin{aligned} \Delta_k &< -\frac{\Delta_{k-1}}{2} - d_0 + \frac{1}{2}\sqrt{\Delta_{k-1}^2 + \frac{4P_r\rho\Delta_{k-1}}{\tau N_0}} && \text{(A.2)} \\ &\approx -\frac{\Delta_{k-1}}{2} - d_0 + \frac{1}{2}\left(\Delta_{k-1} + 2\alpha - \frac{4\alpha^2}{2(\Delta_{k-1} + 2\alpha)}\right) \\ &< \alpha - d_0 < \alpha + d_0 && \text{(A.3)} \end{aligned}$$

where $\alpha = \frac{P_r\rho}{\tau N_o}$. The approximation is obtained by using the identity $\sqrt{z^2 + c} \approx z + \frac{c}{2z}$ and rearranging terms. The upper bound follows since $\Delta_{k-1}, \alpha > 0$.

Substitute the upper bound on Δ_k into the LHS of (A.1):

$$LHS > \frac{(L - 2r_{k-1} - \alpha + \epsilon)^2}{\Delta_k} > \frac{(L - 2\frac{L}{2} - \alpha + \epsilon)^2}{\alpha - d_0} > \alpha$$

where the second inequality follows since $r_{k-1} < \frac{L}{2}$ and the last inequality follows as $\epsilon \to 0$. Hence, we have showed that $\mathrm{SNR}_{k+1}^1(x) > \alpha$, proving our result.

(b) We showed in part (a) that these level sets must lie in the interval $(r_k, L - r_k)$. The SIC receiver decodes signals in the order of decreasing received signal power so

relays in $r_k < x \leq \frac{L}{2}$ will decode symbol s_1 first. In $\frac{L}{2} \leq x < L - r_k$ relays decode symbol s_2 first. Since the levels $\boldsymbol{\mathcal{T}}_k^1$ and $\boldsymbol{\mathcal{T}}_k^2$ are symmetric, we only consider the interval $r_k < x \leq \frac{L}{2}$.

Evaluate the integrals in (2.10) using the values for level sets specified by $\boldsymbol{\mathcal{L}}_k^{(1)}$ to compute $\mathrm{SINR}_{k+1}^1(x)$. We can rewrite $\mathrm{SINR}_{k+1}^1(x) > \tau$ as

$$\theta(x) \geq \tau N_o / P_r \rho (r_{k-1} - r_k) + \psi(x)$$

where $\theta(x) = (x + d_0 - r_k)^{-1}(x + d_0 - r_{k-1})^{-1}$ and $\psi(x) = \tau(L - x + d_0 - r_k)^{-2}$ since $r_k > r_{k-1}$. In the range of interest $(r_k, L - r_k)$, properties (i)-(iii) from Lemma 1 remain valid. Proceeding exactly like the proof of Lemma 1, we get that

$$\begin{aligned} x_k^* \geq x'' &\triangleq \theta^{-1}\left(\frac{\tau N_0}{P_r \rho \Delta_k} + \psi(x')\right) \\ &= \frac{r_k + r_{k-1}}{2} + \frac{1}{2}\sqrt{\Delta_k^2 + \frac{4}{\frac{\tau N_0}{P_r \rho \Delta_k} + \psi(x')}} - d_0. \end{aligned} \tag{A.4}$$

We can verify that $\mathrm{SNR}_{k+1}^2(x) < \tau$, which can be written as

$$(L - x + d_0 - r_{k-1})^{-1}(L - x - r_k + d_0)^{-1} < \tau N_o / P_r \rho \Delta_k,$$

is satisfied for all

$$x < \bar{x}_k = L - \frac{r_k + r_{k-1}}{2} - \frac{1}{2}\sqrt{\Delta_k^2 + \frac{4 P_r \rho \Delta_k}{\tau N_0}} + d_0. \tag{A.5}$$

Finally, one must compare x_k^* and $\bar{x}_k$ in order to determine $\boldsymbol{\mathcal{L}}_{k+1}$. The lemma follows.

(c) The proof is trivial and is omitted.

A.2 Proof of Lemma 4

(a) Consider the right-half of the interval $(q_k, L-q_k)$. Relays in this interval $[\frac{L}{2}, L-q_k)$ will decode symbol s_2 first because the topology of state $\boldsymbol{\mathcal{L}}_k^{(3)}$ dictates that s_2 will be received with greater power. We will solve the system $\mathrm{SINR}_{k+1}^2(x) \geq \tau$ and $\mathrm{SNR}_{k+1}^1(x) < \tau$. By symmetry, the arguments also hold in the left-half $(q_k, \frac{L}{2})$ of the interval.

First consider $\mathrm{SINR}_{k+1}^2(x) \geq \tau$. After taking the limits $N \to \infty$ and $P_r \to 0$, it can be written as $\theta(x) \geq \frac{\tau N_o}{P_r \rho} + \tau\psi(x)$ where $P_r\rho\theta(x)$ and $P_r\rho\psi(x)$ are the received powers corresponding to symbols s_2 and s_1 respectively and

$$\theta(x) = \frac{1}{2}\int_{r_k}^{q_k} \ell(u,x)du + \frac{1}{2}\int_{L-q_k}^{L-r_k} \ell(u,x)du + \int_{L-r_k}^{L-r_{k-1}} \ell(u,x)du,$$

$$\psi(x) = \frac{1}{2}\int_{r_k}^{q_k} \ell(u,x)du + \frac{1}{2}\int_{L-q_k}^{L-r_k} \ell(u,x)du + \int_{r_{k-1}}^{r_k} \ell(u,x)du.$$

Let $\Lambda(x) = \frac{\tau N_o}{P_r \rho} + \tau\psi(x)$. The following statements hold:

(i) $\theta(x)|_{x=\frac{L}{2}} < \Lambda(x)|_{x=\frac{L}{2}}$,

(ii) $\theta'(x) > 0, \forall x \in (\frac{L}{2}, L-q_k+d_0)$ so $\theta(x)$ is monotonically increasing in this interval,

(iii) $\Lambda''(x) > 0, \forall x \in (\frac{L}{2}, L-q_k+d_0)$ so $\Lambda(x)$ is strictly convex in this interval. It follows that $\psi(x)$ is also strictly convex in this interval.

(iv) $\theta(x)$, $\psi(x)$, and $\Lambda(x)$ have an asymptote at $L-q_k+d_0$.

These claims can be verified in a straightforward manner by evaluating the integrals; the details are omitted for brevity.

Properties $(i)-(iv)$ imply that the functions $\theta(x)$ and $\Lambda(x)$ can intersect at most twice in the interval $[\frac{L}{2}, L-q_k+d_0)$ (Fig. A.1). Let $L-q_k+d_0-\delta$ denote the second root of $\theta(x) = \Lambda(x)$, when it exists. If Case C (Fig. A.1) occurs then

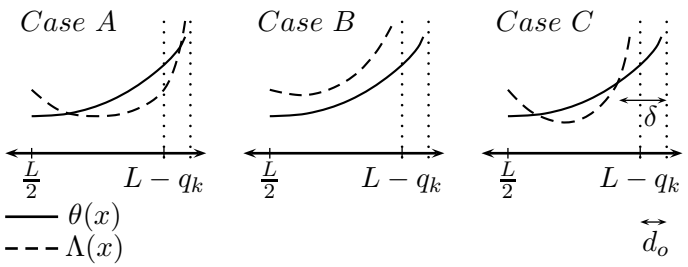


Figure A.1: Based on properties $(i)-(iv)$ defined in Appendix A.2, the functions $\theta(x)$ and $\Lambda(x)$ can intersect at most twice in the interval $[\frac{L}{2}, L - q_k + d_0)$.

statement (i) of Theorem 1 fails since relays within $[L - q_k + d_0 - \delta, L - q_k + d_0)$ decode nothing. Case C cannot happen when $\delta < d_0$, where

$$\delta \approx \frac{1-\tau}{2(\frac{\tau N_o}{P_r \rho} + r_k(\frac{\tau}{(L-q_k+2d_0)(L-q_k-r_k+2d_0)} - \frac{1}{q_k Q_k}))}. \tag{A.6}$$

δ is obtained by solving $\theta(x) = \Lambda(x)$, which is linearized by letting $x = L - q_k + d_0 - \delta$, then neglecting the terms that are small since we are close to the asymptote. We also note that if Case B occurs, then $\text{SINR}_{k+1}^2(x) < \tau$, which means that nothing is decoded.

The decoding region for symbol s_1 is obtained by solving $\text{SNR}_{k+1}^1(x) > \tau \Leftrightarrow \psi(x) > \frac{\tau N_0}{P_r \rho}$, where, from property (iii), $\psi(x)$ is strictly convex. Putting the decoding regions for s_2 and s_1 together yields four possible network states: $\boldsymbol{\mathcal{L}}_{k+1}^{(1)}$ - $\boldsymbol{\mathcal{L}}_{k+1}^{(4)}$.

(b) We will now derive the condition under which the flows cross from state $\boldsymbol{\mathcal{L}}_k^{(3)}$. Consider the interval $(L - r_{k-1}, L]$. We will determine the conditions required such that information flow s_1 will be present in this interval in slot $k+1$. By symmetry, this condition will also indicate the presence of flow s_2 in $[0, r_{k-1})$, thus giving us the condition required to permit flows to cross in slot k from state $\boldsymbol{\mathcal{L}}_k^{(3)}$.

Following the approach used to prove Lemma 1, consider a relay at coordinate

$x > L - r_{k-1}$. Again, because of physical proximity to the level set $\boldsymbol{\mathcal{T}}_k^2$, it will receive s_2 with greater power and decode it first. It is easy to check that there always exists some $x'' \in (L-r_{k-1}, L]$ such that $\text{SINR}_{k+1}^2(x) \geq \tau, \forall x \in (L-r_{k-1}, x'']$. Therefore, we only need to determine the conditions under which $\text{SNR}_{k+1}^1(x) \geq \tau$ for some $x \in (L - r_{k-1}, L]$.

Let $x = L - r_{k-1} + \epsilon$, $\forall \epsilon > 0$ and let $Q_k = q_k - r_k$. Then,

$$\begin{aligned}
\text{SNR}_{k+1}^1(x) &= \frac{P_r \rho}{N_o} \frac{\Delta_k}{(x + d_0 - r_k)(x + d_0 - r_{k-1})} \\
&\quad + \frac{P_r \rho}{N_o} \frac{Q_k}{2(x + d_0 - r_k)(x + d_0 - q_k)} \\
&\quad + \frac{P_r \rho}{N_o} \frac{Q_k}{2(x + d_0 - L + r_k)(x + d_0 - L + q_k)} \qquad \text{(A.7)} \\
&> \frac{P_r \rho}{N_o} \frac{\Delta_k + Q_k}{(L - 2r_{k-1} + d_0 + \epsilon)^2}, \qquad \text{(A.8)}
\end{aligned}$$

where we substituted in our choice of x and used $r_{k-1} < r_k < q_k$. This gives us the required condition

$$P_r \rho (\Delta_k + Q_k) > \tau N_0 (L - 2r_{k-1} + d_0)^2, \qquad \text{(A.9)}$$

as $\epsilon \to 0$. The corresponding derivations for state $\boldsymbol{\mathcal{L}}_{k+1}^{(4)}$ are very similar and are therefore omitted.

A.3 Proof of Lemma 5

Using Theorem 1 statement (ii), any relay $x > L - r_{k-1}$ must have decoded symbol s_2 in a prior slot. We show that under certain conditions, there exists at least one relay in this interval that decodes symbol s_1 in slot $k+1$. Due to symmetry, the same argument shows the existence of a relay $x < r_{k-1}$ that decodes s_2, showing that flows cross.

Recall, $\boldsymbol{\mathcal{T}}_k^b = (r_{k-1}, r_k] \cup [L - r_k, L - r_{k-1})$. Let $\eta = 1$. Then all relays $x > L - r_{k-1}$ must receive s_1 with greater power and decode it first in slot $k+1$ (c.f. Table 2.1). It only remains to find the condition under which they can decode s_1: $\mathrm{SINR}_{k+1}^1(x) \geq \tau$ is equivalent to

$$\int_{L-r_k}^{L-r_{k-1}} \ell(u,x)du \geq \frac{\tau N_o}{P_r\rho} + \tau \int_{r_{k-1}}^{r_k} \ell(u,x)du \tag{A.10}$$

$$\Leftrightarrow \frac{r_k - r_{k-1}}{(x + d_0 - L + r_{k-1})(x + d_0 - L + r_k)} \tag{A.11}$$

$$- \frac{\tau(r_k - r_{k-1})}{(x + d_0 - r_{k-1})(x + d_0 - r_k)} \geq \frac{\tau N_o}{P_r\rho}.$$

Using $r_{k-1} < r_k$, lower bound the second term on the LHS and plug in $x = L - r_{k-1} + \epsilon$, where $\epsilon > 0$, to get

$$\frac{\tau\Delta_k}{(L - 2r_{k-1} + d_0 + \epsilon)^2} \leq \frac{\Delta_k}{(d_0 + \Delta_k + \epsilon)^2} - \frac{\tau N_o}{P_r\rho}. \tag{A.12}$$

We get the desired result as $\epsilon \to 0$.

Proof of Lemma 6: We will derive the necessary conditions under which the information flows can cross from each network state (as given in Theorem 1) under cooperative protocol PCSI (with side information). We consider each network state in turn.

First consider state $\boldsymbol{\mathcal{L}}_k^{(1)}$. Consider relays $x > L - r_{k-1}$. These relays must have decoded symbol s_2 in some previous slot (Theorem 1) and therefore are able to cancel all interference resulting from the transmission of symbol s_2 in future slots. Accordingly, these relays are only interested in decoding s_1, which is transmitted by the elements of the set $\boldsymbol{\mathcal{T}}_k^1 = [r_{k-1}, r_k]$. As a result of the cancellation of interference from s_2, one only needs to check whether $\mathrm{SNR}_{k+1}^1(x) \geq \tau$ for any $x > L - r_k$. Since $\mathrm{SNR}_{k+1}^1(x)$

$$= \frac{P_r\rho}{N_0} \int_{r_{k-1}}^{r_k} \ell(u,x)du = \frac{P_r\rho\Delta_k}{(x + d_0 - r_k)(x + d_0 - r_{k-1})},$$

$\mathrm{SNR}_{k+1}^1(x) \geq \tau$ is a quadratic equation in x with solution

$$x^* < \frac{r_k + r_{k-1}}{2} + \frac{1}{2}\sqrt{\Delta_k^2 + \frac{4P_r\rho\Delta_k}{\tau N_o}}. \tag{A.13}$$

Therefore, since $\boldsymbol{\mathcal{T}}_{k+1}^1$ and $\boldsymbol{\mathcal{T}}_{k+1}^2$ remain symmetric, the information flows cross if $x^* > L - r_{k-1}$.

Now consider the state $\boldsymbol{\mathcal{L}}_k^{(2)}$. Since $\boldsymbol{\mathcal{T}}_k^b \neq \emptyset$, there exist some nodes which forward the symbols of both users using power control. For illustration, let $\eta = 1$. Consider relays $x > L - r_{k-1}$ which have previously decoded s_2 and have side information. Analogous to Lemma 5, since $\mathrm{SNR}_{k+1}^1(x) = \frac{P_r\rho}{N_o}\int_{L-r_k}^{L-r_{k-1}} \ell(u,x)du$, $\mathrm{SNR}_{k+1}^1(x) \geq \tau$ is a quadratic equation whose solution is

$$x^* < L - d_0 - \frac{r_k + r_{k-1}}{2} + \frac{1}{2}\sqrt{\Delta_k^2 + \frac{4P_r\rho\Delta_k}{\tau N_o}}. \tag{A.14}$$

Therefore, the information flows cross if $x^* > L - r_{k-1}$.

Finally, consider the state $\boldsymbol{\mathcal{L}}_k^{(3)}$. Again, consider $x > L - r_{k-1}$ and let $\eta = 1$. Then,

$$\begin{aligned} \mathrm{SNR}_{k+1}^1 &= \frac{P_r\rho}{N_0}\int_{L-q_k}^{L-r_k} \ell(u,x)du + \frac{P_r\rho}{N_0}\int_{r_{k-1}}^{r_k} \ell(u,x)du \\ &\leq (x + d_0 - L + r_k)^{-2} + (x + d_0 - r_{k-1})^{-2} \\ &\leq 2(x + d_0 - r_{k-1})^{-2} \end{aligned} \tag{A.15}$$

since $r_{k-1} < r_k$. $\mathrm{SNR}_{k+1}^1(x) \geq \tau$ then has the solution

$$x^* < r_{k-1} - d_0 + \sqrt{2P_r\rho\Delta_k/\tau N_o}, \tag{A.16}$$

and the information flows cross if $x^* > L - r_{k-1}$. A similar condition can be derived for state $\boldsymbol{\mathcal{L}}_{k+1}^{(4)}$.

A.4 Proof of Lemma 9

The proof draws inspiration from techniques used in [75]. Let $\gamma_n^m \triangleq \text{SINR}_1^m(x_n)$, $\forall\ m$, where x_n denotes the nth relay with abuse of notation. Since the channels $h(u,x) \overset{i.i.d.}{\sim} \mathcal{CN}(0,1)$ are *i.i.d.*, γ_n^m are *i.i.d.* RVs with density and distribution $f_{\gamma_n^m}(\gamma) = e^{-\frac{N_o\tau}{P_s}}(1+\gamma)^{-M}\left(N_o/P_s(1+\gamma)+M-1\right)$ and

$$F_{\gamma_n^m}(\gamma) = 1 - \frac{e^{-\frac{N_o\tau}{P_s}}}{(1+\gamma)^{M-1}},\quad \gamma \geq 0, \tag{A.17}$$

respectively. Let $\gamma_{n_{(r)}}^m$ denote the rth largest realization in the sequence of RVs $\{\gamma_n^m\}_{n=1}^N$, given m. Under the pessimistic scenario of successive decoding *without* interference cancellation, $\gamma_{n_{(r)}}^m \geq \tau$ implies that $N-r$ relays are able to decode s_m [*CLARIFY?*]. This holds for each m. So the question we are asking is: how does $(N-r)$ scale?

Using the distribution of the rth order statistic of *i.i.d.* RVs,

$$P(\gamma_{n_{(r)}}^m \geq \tau) = 1 - \sum_{i=r}^{N} \binom{N}{i} F_{\gamma_n^m}(\tau)^i \left(1 - F_{\gamma_n^m}(\tau)\right)^{N-i}.$$

Note that each summand resembles a Binomial distribution, *i.e.*, $P(Z=i) = \binom{N}{i}p^i(1-p)^{N-i}$ if $Z \sim B(N,p)$ with $p = F_{\gamma_n^m}(\tau)$. It is known that asymptotically, the binomial distribution can be approximated as a gaussian $\frac{Z-Np}{\sqrt{Np(1-p)}} \xrightarrow{d} \mathcal{N}(0,1)$. Using this result we get

$$P(\gamma_{n_{(r)}}^m \geq \tau) \approx 1 - \frac{1}{\sqrt{2\pi Np(1-p)}} \int_r^N e^{-\frac{(i-Np)^2}{2Np(1-p)}} di \tag{A.18}$$

$$= \ 1 - \frac{1}{2}\left[\text{erf}\left(\frac{Np-N}{\sqrt{2Np(1-p)}}\right) - \text{erf}\left(\frac{Np-r}{\sqrt{2Np(1-p)}}\right)\right] \tag{A.19}$$

where $\text{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$. We will show that $P(\gamma_{n_{(r)}}^m \geq \tau) \to 1$. Substituting our choice of $M = \frac{\log(N^{0.5}) - \frac{\tau N_o}{P_s}}{\log(1+\tau)} + 1$ into $F_{\gamma_n^m}(\tau)$, we have that $F_{\gamma_n^m}(\tau) = p = 1 - N^{-\frac{1}{2}}$.

Now,

$$\lim_{N\to\infty} \text{erf}\left[\frac{Np-N}{\sqrt{2Np(1-p)}}\right] = \lim_{N\to\infty} \text{erf}\left[\frac{N-N^{\frac{1}{2}}-N}{\sqrt{2N^{\frac{1}{2}}(1-N^{-\frac{1}{2}})}}\right]$$

$$= \lim_{N\to\infty} \text{erf}\left(-\sqrt{\frac{N^{\frac{1}{2}}}{2(1-N^{-\frac{1}{2}})}}\right) = \text{erf}(-\infty) = -1. \tag{A.20}$$

To complete the proof we need to show that there exists a choice of r such that the right term in (A.19) also tends to -1. Let $N - r = \Theta(\log N)$. Then,

$$\lim_{N\to\infty} \text{erf}\left[\frac{Np-r}{\sqrt{2Np(1-p)}}\right] = \lim_{N\to\infty} \text{erf}\left[\frac{N-N^{\frac{1}{2}}-r}{\sqrt{2N^{\frac{1}{2}}(1-N^{-\frac{1}{2}})}}\right]$$

$$= \lim_{N\to\infty} \text{erf}\left(-\frac{N^{\frac{1}{2}}-\Theta(\log N)}{\sqrt{2N^{\frac{1}{2}}(1-N^{-\frac{1}{2}})}}\right) \tag{A.21}$$

$$= \lim_{N\to\infty} \text{erf}\left(-\frac{O(N^{\frac{1}{2}})}{\sqrt{2N^{\frac{1}{2}}(1-N^{-\frac{1}{2}})}}\right) = -1. \tag{A.22}$$

Thus, we have shown that if $M = O(\log N)$, then each source is able to recruit at least $\Theta(\log N)$ relays w.h.p. to forward its message in the next hop.

A.5 Proof of Lemma 10

We show the derivations for $P_k^1(x)$ and $P_k^b(x)$. The result for $P_k^2(x)$ follows from the expressions obtained for $P_k^1(x)$ because of symmetry between level sets $\boldsymbol{\mathcal{T}}_k^1$ and $\boldsymbol{\mathcal{T}}_k^2$. We drop the node indices from the following for notational convenience since they should be clear from the context. To begin, we note that by plugging in the expression for $\boldsymbol{\mathcal{T}}_i^m$ and using the distributive law of set operations,

$$\boldsymbol{\mathcal{T}}_k^m = (\cup_{A:m\in A} \boldsymbol{\mathcal{S}}_k^A) \cap_{i=1}^{k-1} \bar{\boldsymbol{\mathcal{T}}}_i^m \tag{A.23}$$

$$= (\cup_{A:m\in A} \boldsymbol{\mathcal{S}}_k^A) \cap_{i=1}^{k-1} (\cap_{A:m\in A} \bar{\boldsymbol{\mathcal{S}}}_i^A). \tag{A.24}$$

Then, $P_k^1(x) = P(x \in \boldsymbol{\mathcal{T}}_k^1)$

$$= P\left(x \in \left(\cup_{A:1\in A} \boldsymbol{\mathcal{S}}_k^A\right) \cap_{i=1}^{k-1} \left(\cap_{A:1\in A} \bar{\boldsymbol{\mathcal{S}}}_i^A\right)\right) \tag{A.25}$$

$$\stackrel{(a)}{=} P(x \in \cup_{A:1\in A} \boldsymbol{\mathcal{S}}_k^A) \prod_{i=1}^{k-1} P(x \in \cap_{A:1\in A} \bar{\boldsymbol{\mathcal{S}}}_i^A) \tag{A.26}$$

$$\stackrel{(b)}{=} \sum_{A=\{1\},\{1,2\}} P(x \in \boldsymbol{\mathcal{S}}_k^A) \prod_{i=1}^{k-1} P(x \in \cap_{A=\{1\},\{1,2\}} \bar{\boldsymbol{\mathcal{S}}}_i^A) \tag{A.27}$$

where (a) follows because the aggregate channel gains $\bar{h}_i^m(x)$ are independent over i and (b) follows because the events $\boldsymbol{\mathcal{S}}_k^A$ are mutually exclusive.

Consider the first term of the product in (A.27). We will compute each summand in turn. Let $Z_1 = |\bar{h}_k^1(x)|^2$ and $Z_2 = |\bar{h}_k^1(x)|^2$. From (2.5) it follows that in the continuum limit $N \to \infty$ and $P_r \to 0$, $Z_1 \triangleq |\bar{h}_k^1(x)|^2 \sim \exp(\mu_k^1(x))$ and $Z_2 \triangleq |\bar{h}_k^2(x)|^2 \sim \exp(\mu_k^2(x))$. First, let $A = \{1\}$:

$$\begin{aligned} P(x \in \boldsymbol{\mathcal{S}}_k^1) &= P(x \in \boldsymbol{\mathcal{S}}_k^1,\ Z_1 \geq Z_2) + P(x \in \boldsymbol{\mathcal{S}}_k^1,\ Z_2 \geq Z_1) \\ &= P(x \in \boldsymbol{\mathcal{S}}_k^1,\ Z_1 \geq Z_2) + 0 && \text{(A.28)} \\ &= P(\frac{Z_1}{N_o + Z_2} \geq \tau,\ \frac{Z_2}{N_o} < \tau,\ Z_1 \geq Z_2) && \text{(A.29)} \\ &\triangleq \theta_k(\mu^1, \mu^2), && \text{(A.30)} \end{aligned}$$

where we used the law of total probability to get the first equality. $Z_1 \geq Z_2$ corresponds to the case where the receiver attempts to decode s_1 first. The other equalities follow from the application of the operating rules of the SIC receiver. The event corresponding to $\theta_k(\mu^1, \mu^2)$, when $\tau > 1$, is shown in Fig. A.2.

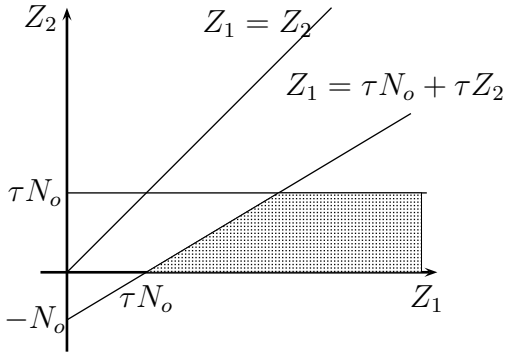


Figure A.2: Probability region: $\theta_k(\mu^1, \mu^2) = P(Z_1 \geq \tau N_o + \tau Z_2,\ Z_2 < \tau N_o,\ Z_1 \geq Z_2)$.

Now consider the remaining terms of the product in (A.27):

$$
\begin{aligned}
P(x \in \bar{\boldsymbol{\mathcal{S}}}_i^1 \cap \bar{\boldsymbol{\mathcal{S}}}_i^b) &= P(x \in \bar{\boldsymbol{\mathcal{S}}}_i^1 \cap \bar{\boldsymbol{\mathcal{S}}}_i^b,\ Z_1 \geq Z_2) \\
&\quad + P(x \in \bar{\boldsymbol{\mathcal{S}}}_i^1 \cap \bar{\boldsymbol{\mathcal{S}}}_i^b,\ Z_2 \geq Z_1) \\
&= P(Z_1 \leq \tau N_o + \tau Z_2,\ Z_1 \geq Z_2) \\
&+ [1 - P(Z_2 \geq \tau N_0 + \tau Z_1,\ Z_1 \geq \tau N_o\ |Z_2 \geq Z_1)]\, P(Z_2 \geq Z_1) \\
&\overset{(a)}{=} [1 - P(Z_2 \geq Z_1) - P(Z_1 \geq \tau N_o + \tau Z_2)] \\
&\quad + P(Z_2 \geq Z_1) - P(\tau N_o \leq Z_1 \leq \frac{Z_2 - \tau N_o}{\tau},\ Z_2 \geq Z_1) \\
&\overset{(b)}{=} 1 - P(Z_1 \geq \tau N_o + \tau Z_2) - P(\tau N_o \leq Z_1 \leq \frac{Z_2 - \tau N_o}{\tau}) \\
&\triangleq 1 - \omega_k(\mu^1, \mu^2) - \gamma_k(\mu^2, \mu^1),
\end{aligned}
\tag{A.31}
$$

where the first term in equality (a) follows by rewriting the probability event to be calculated in terms of its complement set (see Fig. A.2) and the second term in equality (b) follows since the event $Z_2 \geq Z_1$ is implied by the event $\tau N_o \leq Z_1 \leq \frac{Z_2 - \tau N_o}{\tau}$ when $\tau > 1$. Putting all this together gives us the result for $P_k^1(x)$.

Consider $A = \{1, 2\}$, denoted by b. We are left with the task of computing

$P_k^b(x)$:

$$P_k^b(x) = P(x \in \boldsymbol{\mathcal{T}}_k^1 \cap \boldsymbol{\mathcal{T}}_k^2) \tag{A.32}$$

$$= P(x \in \boldsymbol{\mathcal{S}}_k^b) \prod_{i=1}^{k} P(x \in \bar{\boldsymbol{\mathcal{S}}}_i^1 \cap \bar{\boldsymbol{\mathcal{S}}}_i^2 \cap \bar{\boldsymbol{\mathcal{S}}}_i^b). \tag{A.33}$$

Using similar reasoning as before,

$$\begin{aligned}
&P(x \in \boldsymbol{\mathcal{S}}_k^b) = P(x \in \boldsymbol{\mathcal{S}}_k^b,\ Z_1 \geq Z_2) + P(x \in \boldsymbol{\mathcal{S}}_k^b,\ Z_2 \geq Z_1) \\
&= P(Z_1 \geq \tau N_o + \tau Z_2, Z_2 \geq \tau N_0, Z_1 \geq Z_2) \\
&+ P(Z_2 \geq \tau N_o + \tau Z_1, Z_1 \geq \tau N_0, Z_2 \geq Z_1) \\
&= P(\tau N_o \leq Z_2 \leq \frac{Z_1 - \tau N_o}{\tau}) + P(\tau N_o \leq Z_1 \leq \frac{Z_2 - \tau N_o}{\tau}) \\
&\triangleq \gamma_k(\mu^1, \mu^2) + \gamma_k(\mu^2, \mu^1) \\
&= \omega_k(\mu^1, \mu^2) - \theta_k(\mu^1, \mu^2) + \omega_k(\mu^2, \mu^1) - \theta_k(\mu^2, \mu^1).
\end{aligned} \tag{A.34}$$

We can express $\gamma_k(\mu^2, \mu^1)$ as the difference between $\omega_k(\mu^2, \mu^1)$ and $\theta_k(\mu^2, \mu^1)$ by considering the event regions shown in Fig. A.3.

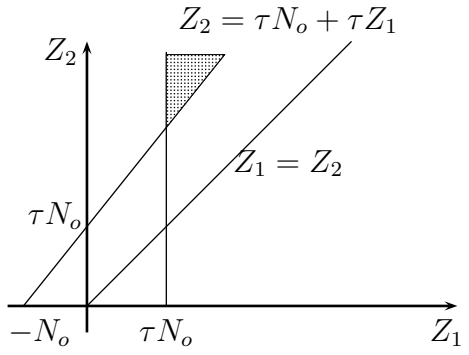


Figure A.3: Probability region: $\gamma_k(\mu^2, \mu^1) = P(\tau N_o \leq Z_1 \leq \frac{Z_2 - \tau N_o}{\tau})$.

Now consider the second term in (A.33):

$$\begin{aligned}
P(x \in \bar{\boldsymbol{\mathcal{S}}}_i^1 \cap \bar{\boldsymbol{\mathcal{S}}}_i^2 \cap \bar{\boldsymbol{\mathcal{S}}}_i^b) &= P(x \in \boldsymbol{\mathcal{S}}_i^{\emptyset},\ Z_1 \geq Z_2) \\
&\quad + P(x \in \boldsymbol{\mathcal{S}}_i^{\emptyset},\ Z_2 \geq Z_1) \\
&= P(\frac{Z_1}{N_o + Z_2} \leq \tau,\ Z_1 \geq Z_2) + P(\frac{Z_2}{N_o + Z_1} \leq \tau,\ Z_2 \geq Z_1) \\
&\stackrel{(a)}{=} [1 - P(Z_2 \geq Z_1) - P(Z_1 \geq \tau N_o + \tau Z_2)] \\
&\quad + [1 - P(Z_1 \geq Z_2) - P(Z_2 \geq \tau N_o + \tau Z_1)] \\
&= 1 - P(Z_1 \geq \tau N_o + \tau Z_2) - P(Z_2 \geq \tau N_o + \tau Z_1) \\
&= 1 - \omega_k(\mu^1, \mu^2) - \omega_k(\mu^2, \mu^1) \qquad \text{(A.35)}
\end{aligned}$$

where equality (a) is obtained by rewriting the probability region to be calculated in terms of its complement set (see Figs. A.2 and A.3), and using $P(Z_2 \geq Z_1) + P(Z_1 \geq Z_2) = 1$.

Now all we need to do is compute the probabilities $\theta_k(\mu^1, \mu^2)$ and $\omega_k(\mu^1, \mu^2)$. Since $Z_1 \perp Z_2$, the calculations are straightforward and the result as stated in Lemma 10 can be obtained.

A.6 Proof of Lemma 12

No side-information is available initially, so $\bar{P}_1^i(x) = P_1^i(x)$, $i = 1, 2, b$. Since $x \in \boldsymbol{\mathcal{T}}_{k+1}^b$ can only occur if x has no prior side-information, $\bar{P}_{k+1}^b(x)$ is unchanged from protocol PC. Define $\xi^m(x) = \mathbf{1}_{\{\text{already decoded } s_m\}}(x)$ where $\mathbf{1}_{\{\cdot\}}(x)$ denotes the indicator function. Furthermore, if the current slot is $k+1$, we know that $P(\xi^m(x) = 1) = \sum_{j=1}^k \bar{P}_j^m(x)$ using the independence of the aggregate channel

gains over slots. Then, using the law of total probability,

$$\begin{aligned}\bar{P}^m_{k+1}(x) &= P(x \in \boldsymbol{\mathcal{T}}^m_{k+1} | \xi^m(x) = 1) P(\xi^m(x) = 1) \\ &\quad + P(x \in \boldsymbol{\mathcal{T}}^m_{k+1} | \xi^m(x) = 0) P(\xi^m(x) = 0) \\ &\overset{(a)}{=} P(|\bar{h}^m_k(x)|^2 \geq \tau N_o) \prod_{i=1}^{k} P(x \in \cap_{A: m \in A} \bar{\boldsymbol{\mathcal{S}}}^A_i) \\ &\quad \times P(\xi^m(x) = 1) \; + \; P^m_{k+1}(x) P(\xi^m(x) = 0).\end{aligned}$$

where (a) holds since x can cancel the interference from user $n \in \{1, 2\}$, $m \neq n$. The final result is obtained by noting that $P(|\bar{h}^m_k(x)|^2 \geq \tau N_o) = e^{-\tau N_o \mu^m_k(x)}$ and $P(x \in \cap_{A: m \in A} \bar{\boldsymbol{\mathcal{S}}}^A_i) = (1 - P^m_i(x))$.

APPENDIX B

PROOFS FOR CHAPTER 3

B.1 Proof of Lemma 15

The proof derives the joint characteristic function of the random variables $\lim_{\substack{N\to\infty \\ P_r\to 0}}(z_j^{(k+1)}, v_j^{(k+1)})$. In the proof the function $\Phi_X(f)$ will be always the characteristic function of the random variable X, i.e. $\Phi_X(f) = \mathbb{E}\{e^{-j2\pi fX}\}$:

$$\begin{aligned}
\Phi(f_1, f_2) &= \mathbb{E}\left\{ \lim_{\substack{N\to\infty \\ P_r\to 0}} e^{-j2\pi(f_1 z_j^{(k+1)} + f_2 v_j^{(k+1)})} \right\} \\
&= \mathbb{E}\left\{ \lim_{\substack{N\to\infty \\ P_r\to 0}} \prod_{i=1}^{N} \Phi_{e^{j\theta_{ji}}}\left(\sqrt{\frac{\bar{P}_r \alpha_{ji} \ell(d_{ji})}{N}} \epsilon_i^{(k)} (f_1 + f_2 e_i^{(k)}) \right) \right\} \\
&= \mathbb{E}\left\{ \lim_{\substack{N\to\infty \\ P_r\to 0}} \prod_{i=1}^{N} \left(1 - \frac{4\pi^2 \bar{P}_r \ell(d_{ji}) \alpha_{ji} \epsilon_i^{2(k)} (f_1 + f_2 e_i^{(k)})^2}{N} + o(\frac{1}{N}) \right) \right\} \\
&= \mathbb{E}\left\{ \lim_{N\to\infty} e^{-\frac{1}{N}\sum_{i=1}^{N} 4\pi^2 \bar{P}_r \ell(d_{ji}) \alpha_{ji} \epsilon_i^{2(k)} (f_1 + f_2 e_i^{(k)})^2} \right\} \qquad \text{(B.1)}
\end{aligned}$$

Therefore if we take the expectation of the term inside the exponent, as argued before we can show that:

$$\begin{aligned}
M_N &= \mathbb{E}\left\{ \frac{1}{N} \sum_{i=1}^{N} 4\pi^2 \bar{P}_r \ell(d_{ji}) \alpha_{ji} \epsilon_i^{2(k)} (f_1 + f_2 e_i^{(k)})^2 \right\} \\
\lim_{N\to\infty} M_N &= 4\pi^2 \bar{P}_r \iint_{L_k} \ell(x-u, y-v) \left(f_1^2 + \right. \\
&\quad \left. + 2 f_1 f_2 \psi^{(k)}(u,v) + f_2^2 \psi^{(k)}(u,v) \right) du dv.
\end{aligned}$$

Let us call X_N the random variable:

$$X_N = \frac{1}{N} \sum_{i=1}^{N} 4\pi^2 \bar{P}_r \ell(d_{ji}) \alpha_{ji} \epsilon_i^{2(k)} (f_1 + f_2 e_i^{(k)})^2,$$

which for the Law of Large Numbers is such that, for an arbitrarily small $\delta > 0$:

$$\lim_{N\to\infty} P(|M_N - X_N| > \delta) = 0$$

Clearly: $\Phi(f_1, f_2) = e^{-\lim_{N\to\infty} M_N} - o(\delta)$. The expression of $\Phi(f_1, f_2)$ is that of the characteristic function of two jointly Gaussian random variables and by inspection it can be verified that the parameters of the quadratic form $\lim_{N\to\infty} M_N$ give the variances and correlation coefficients in Lemma 15.

B.2 Proof of Corollary 3

In the following proof we drop the sub/super-scripts in expressions where there is no ambiguity for notational convenience. From equation (3.23),

$$\begin{aligned}
\mathbb{E}\left\{P\left(e_j^{(k)}=1\right)\right\} &= \psi^{(k)}(x,y) \qquad \text{(B.2)}\\
&\leq \frac{1}{2}\mathbb{E}\left\{e^{-\frac{\left|z_{(x,y)}^{(k)}\right|^2}{N_o}\left(1-\frac{2\Re\left\{z_{(x,y)}^{*(k)}v_{(x,y)}^{(k)}\right\}}{\left|z_{(x,y)}^{(k)}\right|^2}\right)^2}\right\}\\
&= \frac{1}{2}\mathbb{E}_z\left\{\mathbb{E}_{v|z}\left\{e^{-\frac{\left|z_{(x,y)}^{(k)}\right|^2}{N_o}\left(1-\frac{2\Re\left\{z_{(x,y)}^{*(k)}v_{(x,y)}^{(k)}\right\}}{\left|z_{(x,y)}^{(k)}\right|^2}\right)^2}\right\}\right\}
\end{aligned}$$

Let $s = -\frac{\left|z_{(x,y)}^{(k)}\right|^2}{N_o}$ and $t = \left(1 - \frac{2\Re\left\{z_{(x,y)}^{*(k)}v_{(x,y)}^{(k)}\right\}}{\left|z_{(x,y)}^{(k)}\right|^2}\right)$. By utilizing the Gaussian probability density function of $\lim_{\substack{N\to\infty\\P_r\to 0}}(z_j^{(k+1)}, v_j^{(k+1)})$ found in lemma 15, it follows that the conditional distribution of t, given that $z_{(x,y)}^{(k)}$ is known, is $t \sim \mathcal{CN}(\mu, \sigma^2)$ where $\mu = \left(1 - 2\frac{\nu_{(x,y)}^{(k)}}{\xi_{(x,y)}^{(k)}}\right)$ and $\sigma^2 = \frac{2}{\left|z_{(x,y)}^{(k)}\right|^2}\nu_{(x,y)}^{(k)}\left[1 - \frac{\nu_{(x,y)}^{(k)}}{\xi_{(x,y)}^{(k)}}\right]$. It is easily shown that

$$\mathbb{E}\left\{e^{st^2}\right\} = \frac{1}{\sqrt{1-2s\sigma^2}}e^{\frac{s\mu^2}{1-2s\sigma^2}}.$$

By substituting these expressions into equation (B.2) we thus get,

$$\begin{aligned}
\psi^{(k)}(x,y) &\leq \frac{1}{2}\mathbb{E}_{z^{(k)}_{(x,y)}}\left\{e^{st^2}\right\} \\
&= \frac{1}{2}\mathbb{E}_z\left\{\frac{1}{\sqrt{1+\frac{4}{N_o}\nu\left(1-\frac{\nu}{\xi}\right)}}e^{\frac{-\frac{|z|^2}{N_o}\left(1-2\frac{\nu}{\xi}\right)^2}{1+\frac{4}{N_o}\nu\left(1-\frac{\nu}{\xi}\right)}}\right\} \\
&= \frac{1}{2}\frac{1}{\sqrt{1+\frac{4}{N_o}\nu\left(1-\frac{\nu}{\xi}\right)}}\mathbb{E}_z\left\{e^{-|z|^2G}\right\} \\
&= \frac{1}{2}\frac{1}{\sqrt{1+\frac{4}{N_o}\nu\left(1-\frac{\nu}{\xi}\right)}}\cdot\frac{1}{1+G\xi} \qquad \text{(B.3)}
\end{aligned}$$

where we have used that $\left|z^{(k)}_{(x,y)}\right|^2 \sim exp\left(\frac{1}{\xi^{(k)}_{(x,y)}}\right)$ and

$$G^{(k)}_{(x,y)} = \frac{\frac{1}{N_o}\left(1-2\frac{\nu^{(k)}_{(x,y)}}{\xi^{(k)}_{(x,y)}}\right)^2}{1+\frac{4}{N_o}\nu^{(k)}_{(x,y)}\left(1-\frac{\nu^{(k)}_{(x,y)}}{\xi^{(k)}_{(x,y)}}\right)} \qquad \text{(B.4)}$$

B.3 Proof of Corollary 4

After plugging-in the expression for $G^k(x,y)$, equation (3.28) [c.f. Corollary 3] can be rearranged as follows:

$$\psi^k(x,y) \leq \frac{0.5}{\sqrt{1+\frac{\xi}{N0}}}\cdot\sqrt{1-\left(1-\frac{\nu}{\xi}\right)^2+\frac{1}{1+\frac{\xi}{N0}}\left(1-\frac{\nu}{\xi}\right)^2} \qquad \text{(B.5)}$$

We know that for any given level the following inequalities are true:

$$0 \leq \frac{\nu^k(x,y)}{\xi^k(x,y)} \leq \lambda_k \leq 0.5 \Rightarrow \left(1-\frac{2\nu^k(x,y)}{\xi^k(x,y)}\right) \leq 1, \qquad \text{(a)}$$

$$1-\left(1-\frac{2\nu^k(x,y)}{\xi^k(x,y)}\right)^2 \leq 1-(1-2\lambda_k)^2, \qquad \text{(b)}$$

$$\frac{1}{1+\frac{\xi^k(x,y)}{N0}} \leq \frac{1}{1+\frac{\alpha^2}{2}}, \qquad \text{(c)}$$

and from (a) and (c),

$$\frac{1}{1+\frac{\xi^k(x,y)}{N0}}\left(1-\frac{2\nu^k(x,y)}{\xi^k(x,y)}\right)^2 \leq \frac{1}{1+\frac{\alpha^2}{2}} \tag{d}$$

By applying (a)-(d) to equation (B.5) we get:

$$\begin{aligned}\psi^k(x,y) &< \lambda_k \\ &\leq \frac{0.5}{\sqrt{1+\frac{\alpha^2}{2}}}\sqrt{1-(1-2\lambda_{k-1})^2+\frac{1}{1+\frac{\alpha^2}{2}}}\end{aligned} \tag{B.6}$$

By squaring both sides of the above we obtain a second order equation:

$$\lambda^2\left(2+\frac{\alpha^2}{2}\right)-\lambda-\frac{1}{4\left(1+\frac{\alpha^2}{2}\right)}=0 \tag{B.7}$$

This equation has a real positive and a real negative root. Since λ cannot be negative, the only solution is as indicated in Corollary 4. The solution for the fixed point exceeds 0.5 if $\alpha^2/2 < 0.618$. Hence, in this regime the bound becomes too loose to be meaningful.

APPENDIX C

PROOFS FOR CHAPTER 4

C.1 Proof sketch for Theorem 2

First we bound ϕ_k. By conditioning on $|h_{ij}|^2$, from (4.3) we have $\sqrt{\phi_t} = P(SIR_{(i,j)\in\mathcal{S}_t} \geq \tau) = \mathbb{E}_{|h_{ij}|^2} P(|h_{ij}|^2 \geq \tau \sum_{\ell\in\mathcal{B}_t,\ell\neq i} |h_{\ell j}|^2)$, where $|h_{\ell j}|^2$ are *inid* exponential RVs. We lower bound ϕ_t, clarifying the dependence on k, as follows:

$$\sqrt{\phi_t} \geq \mathbb{E}_{|h_{ij}|^2} P\left(\sum_{\ell=1}^{Q-1} Z_\ell \leq |h_{ij}|^2/\tau\right), \tag{C.1}$$

where $Z_\ell = |h_{mj}|^2$, $\forall \ell$. Index m is chosen such that it causes the most interference to node j on average, *i.e.* its closest to node j. So $m = \arg\min_{m\in\mathcal{B}_t,m\neq i} d_{mj}$ with $d_{mj} = (2k+1)d_o - kd_o = (k+1)d_o$ where $d_o \triangleq d_{i,i+1}$. So $Z_\ell \overset{iid}{\sim} \exp(1/\alpha_{mj})$ and $\sum_{\ell=1}^{Q-1} Z_\ell \sim G(Q-1, 1/\alpha_{mj})$ has a gamma density. Using the gamma distribution and taking the expectation with respect to $|h_{ij}|^2$ we obtain

$$\sqrt{\phi_t} \geq 1 - \sum_{\ell=0}^{Q-2} \frac{\alpha_{mj}}{\left(\frac{\alpha_{mj}}{\alpha_{ij}} + \frac{1}{\tau}\right)^{\ell+1}} \geq 1 - \frac{(Q-1)\alpha_{mj}}{\left(\frac{\alpha_{mj}}{\alpha_{ij}} + \frac{1}{\tau}\right)}, \tag{C.2}$$

where $\alpha_{mj} = (1+(1+k)d_o)^{-\gamma} < \alpha_{ij} = (1+kd_o)^{-\gamma}$. The second inequality is obtained by noting that $\frac{\alpha_{mj}}{\alpha_{ij}} + \frac{1}{\tau} \geq 1$ when τ is small since $\frac{\alpha_{mj}}{\alpha_{ij}}$ is close to 1. Substituting the values of α_{mj} and α_{ij} into (C.2), when $\frac{N(k-1)}{Q} \leq t \leq \frac{Nk}{Q} - 1, k \in [1, K]$,

$$\sqrt{\phi_k} \geq 1 - \frac{(Q-1)(1+(1+k)d_o)^{-\gamma}}{\frac{1}{\tau} + \frac{(1+kd_o)^\gamma}{(1+(1+k)d_o)^\gamma}} = 1 - O\left(\frac{1}{k^{\gamma+1}}\right),$$

since $Q = \frac{N}{2k+1}$. This proves the first part of the theorem.

Now we bound $\lambda_2(\mathbb{E}\{W^T(T)W(T)\})$. The key step is showing that $\mathbb{E}\{W^T(T)W(T)\}$, can be approximated by a diagonalizable circulant matrix.

Plugging-in (4.4) into $W^T(T)W(T) = \left(\prod_{t=T-1}^{0} W_{\mathcal{S}_t}^T\right)\left(\prod_{t=0}^{T-1} W_{\mathcal{S}_t}\right)$, and expanding it, we get $W^T(T)W(T) = I - \frac{1}{2}\sum_{(i,j)\in\{\mathcal{S}_0,\ldots,\mathcal{S}_{T-1}\}}\Big[(\mathbf{e}_i-\mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^T\mathbb{1}_{\{\mathrm{SIR}_{(i,j)},\mathrm{SIR}_{(j,i)}\geq\tau\}}$ $+ (\mathbf{e}_i-\mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^T\mathbb{1}_{\{\mathrm{SIR}_{(i,j)},\mathrm{SIR}_{(j,i)}\geq\tau\}}\Big] + \mathrm{cross-terms}$, where the last equality holds because $W_{\mathcal{S}_t}$ is symmetric. Taking the expectation on both sides, we see that each cross-term is pre-multiplied by $\phi_t^z, z > 1$, and can be neglected since $\phi_t < 1, \forall\ t$. So, $\mathbb{E}\{W(T)^TW(T)\} \approx I - \sum_{(i,j)\in\{\mathcal{S}_0,\ldots,\mathcal{S}_{T-1}\}}(\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^T\phi_t$. Decomposing the first T slots into chunks of $\frac{N}{Q}$ slots we get, $\mathbb{E}\{W(T)^TW(T)\} \approx I - \sum_{k=1}^{K}\phi_k\sum_{(i,j)\in\{\mathcal{S}_{\frac{N(k-1)}{Q}},\ldots,\mathcal{S}_{\frac{Nk}{Q}-1}\}}(\mathbf{e}_i-\mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^T$, where $\phi_k\sum_{(i,j)\in\{\mathcal{S}_{\frac{N(k-1)}{Q}},\ldots,\mathcal{S}_{\frac{Nk}{Q}-1}\}}(\mathbf{e}_i-\mathbf{e}_j)(\mathbf{e}_i-\mathbf{e}_j)^T$, which we denote as A_k, is circulant: $A_k = \mathrm{circ}(2\phi_k, \mathbf{0}_{k-1}, -\phi_k, \mathbf{0}_{N-2k-1}, -\phi_k, \mathbf{0}_{k-1})$, where $\mathbf{0}_k$ is a $1\times k$ vector of all zeros.

An $N\times N$ circulant matrix can be diagonalized by the N-point DFT matrix F_N [20]. So $F_NA_kF_N^H$ is diagonal with entries

$$\lambda_{s+1}^k(A_k) = 2\phi_k(1-\cos\frac{2\pi ks}{N}),\ s=0,\ldots,N-1, \tag{C.3}$$

which are the eigenvalues of A_k. Moreover, the sum of circulant matrices remains circulant. Therefore, $\mathbb{E}\{W(T)^TW(T)\}$ can also be diagonalized by F_N and its eigenvalues are $\lambda_{s+1}(\mathbb{E}\{W(T)^TW(T)\}) \approx 1 - \sum_{k=1}^{K}2\phi_k(1-\cos\frac{2\pi ks}{N}),\ s = 0,\ldots,N-1$. $s=0$ yields the largest eigenvalue $\lambda_1(\mathbb{E}\{W(T)^TW(T)\}) = 1$. The second largest eigenvalue is given by $s=1$ or $N-1$, so

$$\lambda_2(\mathbb{E}\{W(T)^TW(T)\}) \approx 1 - 4\sum_{k=1}^{K}\phi_k\sin^2\tfrac{\pi k}{N} \tag{C.4}$$

$$\leq\ 1 - 4\phi_K\sin^2(\pi K/N) \tag{C.5}$$

$$=\ 1 - 4\sin^2(\pi K/N)\left(1 - O\left(K^{-\gamma-1}\right)\right). \tag{C.6}$$

We used $1-\cos x = 2\sin^2\frac{x}{2}$ to get (C.4). Since all the summands in (C.4) are positive, we drop the first $K-1$ summands to get (C.5). Finally, we use $\phi_K \geq$

$(1 - O(K^{-\gamma-1}))^2 = 1 + O(K^{-2(\gamma+1)}) - O(K^{-\gamma-1}) \geq 1 - O(K^{-\gamma-1})$ since $\gamma > 0$ to get the desired result.

APPENDIX D

PROOFS FOR CHAPTER 5

D.1 Proof of Lemma 21.

Proof Recall the vector form of the data driven update equation. Then,

$$\begin{aligned} \mathbf{x}(t+1) &= \bar{\mathbf{x}}(t) + \hat{\mathbf{u}}(t) \\ &= \bar{\mathbf{x}}(t) + \mathbb{E}\{\hat{\mathbf{u}}(t)\} + \hat{\mathbf{u}}(t) - \mathbb{E}\{\hat{\mathbf{u}}(t)\} \quad \text{(D.1)} \\ &= W\bar{\mathbf{x}}(t) + \mathbf{e}(t) \quad \text{(D.2)} \end{aligned}$$

where $\mathbf{e}(t) \triangleq \hat{\mathbf{u}}(t) - \mathbb{E}\{\hat{\mathbf{u}}(t)\}$ and the last equality follows from Lemma 19 which showed that the average behavior of data driven consensus is the same as standard average consensus using quantized states. Writing the quantized state vector $\bar{\mathbf{x}}(t)$ in terms of the quantization error $\mathbf{v}(t)$, whose statistics were defined in equations (5.7)-(5.8), and expanding the above recursion we get

$$\mathbf{x}(t) = W^t\mathbf{x}(0) + \sum_{i=1}^{t} W^i\mathbf{v}(t-i) + \sum_{i=1}^{t} W^{i-1}\mathbf{e}(t-i). \quad \text{(D.3)}$$

Recall that $\mathbf{v}(t_1)$ and $\mathbf{v}(t_2)$, $t_1 \neq t_2$ are uncorrelated. Given $\mathbf{x}(0)$, so are $\mathbf{e}(t_1)$ and $\mathbf{e}(t_2)$. Using the above expression for $\mathbf{x}(t)$ in the definition of MSE, with some algebra we get

$$\begin{aligned} MSE(t) &= \frac{1}{N}\mathbb{E}\{||(W^t - J)\mathbf{x}(0)||^2\} + \frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^i\mathbf{v}(t-i)||^2\} \quad \text{(D.4)} \\ &\quad + \frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^{i-1}\mathbf{e}(t-i)||^2\} \quad \text{(D.5)} \end{aligned}$$

This reveals that the average MSE comprises three contributions - convergence error, quantization error, and channel error respectively. We know the first term

through Lemma 16. Let us consider the quantization error term in (D.4). Taking the trace of the scalar norm and reordering its argument, we get

$$\begin{aligned}
\frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^i\mathbf{v}(t-i)||^2\} &= \frac{1}{N}\sum_{i=t}^{1} Tr\left(W^i\mathbb{E}\{\mathbf{v}(t-i)\mathbf{v}^T(t-i)\}W^i\right) \\
&= \frac{1}{N}\sum_{i=t}^{1} Tr\left(\mathbb{E}\{\mathbf{v}(t-i)\mathbf{v}^T(t-i)\}W^{2i}\right) && \text{(D.6)} \\
&= \frac{1}{N}\frac{\Delta^2}{12}\sum_{i=t}^{1}\sum_{j=1}^{N}\lambda_j^{2i}(W) && \text{(D.7)} \\
&= f(\bar{\Delta})\frac{1}{N}\sum_{i=t}^{1}\left(1+\sum_{j=2}^{N}\lambda_j^{2i}(W)\right) && \text{(D.8)} \\
&= f(\bar{\Delta})\left(\frac{t}{N}+\frac{1}{N}\sum_{j=2}^{N}N\left(\frac{1-\lambda_j^{2(t+1)}(W)}{1-\lambda_j^2(W)}-1\right)\right) && \text{(D.9)} \\
&\leq f(\bar{\Delta})\left(\frac{t}{N}+\max_{2\leq j\leq N}\frac{\lambda_j^2(W)(1-\lambda_j^{2t})(W)}{1-\lambda_j^2(W)}\right) && \text{(D.10)} \\
&= f(\bar{\Delta})\left(\frac{t}{N}+\frac{\lambda_2^2(W)(1-\lambda_2^{2t}(W))}{1-\lambda_2^2(W)}\right) && \text{(D.11)}
\end{aligned}$$

where $\mathbb{E}\{\mathbf{v}(t-i)\mathbf{v}^T(t-i)\} = \frac{C^2\sigma^2}{12Q^2}I \triangleq \frac{\Delta^2}{12}I$ gives (D.7). Noting that $\lambda_1(W) = 1$ and defining $f(\bar{\Delta}) = \frac{\Delta^2}{12}$ gives (D.8). Computing the finite geometric sum over t gives (D.9). For any $t \geq 0$, the second largest eigenvalue of W, $\lambda_2(W)$ with $\lambda_2^2(W) \in (0,1)$, maximizes the term in (D.10). Finally, note that

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^i\mathbf{v}(t-i)||^2\} = f(\bar{\Delta})\left(\frac{\lambda_2^2(W)(1-\lambda_2^{2t}(W))}{1-\lambda_2^2(W)}\right) = O(f(\bar{\Delta})). \tag{D.12}$$

Now consider the last error term (D.5). Recall that by definition, $\mathbb{E}\{\mathbf{e}(t)\} = 0, \forall t$, and $[\mathbb{E}\{\mathbf{e}(t-i)\mathbf{e}^T(t-i)\}]_{j,j} = VAR(\hat{\mathbf{u}}_j(t-i))$. Proceeding like before,

$$\begin{aligned}
\frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^{i-1}\mathbf{e}(t-i)||^2\} &= \frac{1}{N}\sum_{i=t}^{1} Tr\left(\mathbb{E}\{\mathbf{e}(t-i)\mathbf{e}^T(t-i)\}\}W^{2(i-1)}\right) \\
&= \frac{1}{N}\sum_{i=t}^{1}\sum_{j=1}^{n} VAR(\hat{\mathbf{u}}_j(t-i))\lambda_j^{2(i-1)}(W). && \text{(D.13)}
\end{aligned}$$

Now,

$$VAR(\hat{\mathbf{u}}_j(t)|\bar{\mathbf{x}}(0)) = \frac{1}{\Psi}\sum_{\ell=1}^{Q}(q_\ell - \bar{\mathbf{x}}_j(t))^2 VAR(|r_j(t,\ell)|^2) \tag{D.14}$$

$$= \frac{1}{\Psi}\sum_{\ell=1}^{Q}(q_\ell - \bar{\mathbf{x}}_j(t))^2 \left[\sum_{s=1}^{N}\frac{\alpha_{js}}{N}\delta[q_\ell - \bar{\mathbf{x}}_s(t)] + \frac{N_0}{N}\right]^2 \tag{D.15}$$

$$= \frac{1}{\Psi}\sum_{\ell=1}^{Q}(q_\ell - \bar{\mathbf{x}}_j(t))^2 \left(\frac{N_0}{N}\right)^2 \tag{D.16}$$

$$+\frac{2}{\Psi}\left(\frac{N_0}{N}\right)\sum_{s=1}^{N}\frac{\alpha_{js}^2}{N^2}(\bar{\mathbf{x}}_s(t) - \bar{\mathbf{x}}_j(t))^2$$

$$+\frac{1}{\Psi}\sum_{s=1}^{N}\frac{\alpha_{js}}{N}(\bar{\mathbf{x}}_s(t) - \bar{\mathbf{x}}_j(t))^2 \sum_{s'=1}^{N}\frac{\alpha_{js'}}{N} I_{[\bar{\mathbf{x}}_{s'}(t)=\bar{\mathbf{x}}_j(t)]}.$$

where Ψ is the order of the repetition code. Let us define $\boldsymbol{\beta}(t) \triangleq \bar{\mathbf{x}}(t) - J\mathbf{x}(t)$. We simplify the last term above as

$$\frac{1}{\Psi}\sum_{s=1}^{N}\frac{\alpha_{js}}{N}(\bar{\mathbf{x}}_s(t) - \bar{\mathbf{x}}_j(t))^2 \sum_{s'=1}^{N}\frac{\alpha_{js'}}{N} I_{[\bar{\mathbf{x}}_{s'}(t)=\bar{\mathbf{x}}_j(t)]} \leq \frac{\xi_j}{\Psi N}\sum_{s=1}^{N}(\bar{\mathbf{x}}_s(t) - \bar{\mathbf{x}}_j(t))^2 \tag{D.17}$$

$$= \frac{\xi_j}{\Psi N}\sum_{s=1}^{N}\left(\boldsymbol{\beta}_s(t) - \boldsymbol{\beta}_j(t)\right)$$

$$= \frac{\xi_j}{\Psi N}||\boldsymbol{\beta}(t) - \boldsymbol{\beta}_j(t)\mathbf{1}||^2$$

$$= \frac{\xi_j}{\Psi N}\left(||\boldsymbol{\beta}(t)||^2 + N|\boldsymbol{\beta}_j(t)|^2\right),$$

where ξ_j is some constant that upper bounds $(\max_s \alpha_{js})\sum_{s'=1}^{N}\frac{\alpha_{js'}}{N} I_{[\bar{\mathbf{x}}_{s'}(t)=\bar{\mathbf{x}}_j(t)]}$. The second equality follows by adding and subtracting $J\mathbf{x}(t)$ and using the definition of $\boldsymbol{\beta}(t)$. In the last equality the cross-term $2|\boldsymbol{\beta}_j(t)|\boldsymbol{\beta}(t)^T\mathbf{1}$ can be neglected because $\boldsymbol{\beta}(t)^T\mathbf{1} = \bar{\mathbf{x}}(t)^T\mathbf{1} - \mathbf{x}(t)^T J\mathbf{1} = \mathbf{v}(t)^T\mathbf{1}$ using $J\mathbf{1} = \mathbf{1}$, and $\lim_{N\to\infty}\mathbf{v}(t)^T\mathbf{1} = \lim_{N\to\infty}\sum_{j=1}^{N}\mathbf{v}_j(t) = \mathbb{E}\{\mathbf{v}_j(t)\} = 0$ by WLLN. By definition, $\alpha_{js'} < 1,\ \forall j, s'$, so $\xi_j < 1$. Thus, we can upper bound the variance as,

$$VAR(\hat{\mathbf{u}}_j(t)) \leq O\left(\frac{N_0^2}{\Psi N^2}\right) + O\left(\frac{N_0}{\Psi N}\right) + \frac{\xi_j}{\Psi N}\mathbb{E}\{||\boldsymbol{\beta}(t)||^2 + N|\boldsymbol{\beta}_j(t)|^2\} \tag{D.18}$$

where the first two terms vanish either when the SNR is high, N is large, or a high order repetition code is used. Therefore,

$$\begin{aligned}
&\frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^{i-1}\mathbf{e}(t-i)||^2\} \qquad (D.19)\\
&\leq \frac{1}{N}\frac{1}{\Psi}\sum_{i=t}^{1}\sum_{j=1}^{N}\xi_j\left[\frac{1}{N}\mathbb{E}\{||\boldsymbol{\beta}(t-i)||^2\}+\mathbb{E}\{|\boldsymbol{\beta}_j(t-i)|^2\}\right]\lambda_j^{2(i-1)}(W)\\
&\quad +O\left(\frac{N_0}{\Psi N}\right)+O\left(\frac{N_0^2}{\Psi N^2}\right)\\
&\leq \frac{2\xi}{N}\frac{1}{\Psi}\sum_{i=t}^{1}\lambda_2^{2(i-1)}(W)\mathbb{E}\{||\boldsymbol{\beta}(t-i)||^2\}+O\left(\frac{1}{N}\right)+O\left(\frac{N_0}{\Psi N}\right)+O\left(\frac{N_0^2}{\Psi N^2}\right),
\end{aligned}$$

where $\xi=\max_j\xi_j$. Now, using the recursion (D.3), $\boldsymbol{\beta}(k)$ can be written as

$$\begin{aligned}
\boldsymbol{\beta}(k) &= \bar{\mathbf{x}}(k)-J\mathbf{x}(k)=(I-J)\mathbf{x}(k)+\mathbf{v}(k) \qquad (D.20)\\
&= (W^t-J)\mathbf{x}(0)+\sum_{\ell=k}^{1}(W-J)^{\ell}\mathbf{v}(k-l)+\mathbf{v}(k)+\sum_{\ell=k}^{1}(W-J)^{\ell-1}\mathbf{e}(k-\ell).
\end{aligned}$$

Noting that $W^kJ=J,\forall k$, with some algebra it can be seen that

$$\frac{1}{N}\mathbb{E}\{||\boldsymbol{\beta}(k)||^2\}\leq MSE(k)+\frac{\mathbb{E}\{||\mathbf{v}(k)||^2\}}{N} \qquad (D.21)$$

where $\frac{\mathbb{E}\{||\mathbf{v}(k)||^2\}}{N}=\frac{\Delta^2}{12}=f(\bar{\Delta})$. Therefore, we finally get

$$\begin{aligned}
\frac{1}{N}\sum_{i=t}^{1}\mathbb{E}\{||W^{i-1}\mathbf{e}(t-i)||^2\} &\leq \frac{2\xi}{\Psi}\sum_{i=t}^{1}\lambda_2^{2(i-1)}(W)MSE(t-i)+O(f(\bar{\Delta}))\\
&\quad +O\left(\frac{1}{N}\right)+O\left(\frac{N_0}{\Psi N}\right)+O\left(\frac{N_0^2}{\Psi N^2}\right) \qquad (D.22)
\end{aligned}$$

since $\frac{2\xi}{\Psi}\sum_{i=t}^{1}\lambda_2^{2(i-1)}(W)$ converges to some constant.

Putting everything together and taking the limit $N\to\infty$, we get

$$\begin{aligned}
\lim_{N\to\infty}MSE(t) &\leq \lambda_2^{2t}\lim_{N\to\infty}\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}+O(f(\bar{\Delta}))+\frac{2\xi}{\Psi}\sum_{i=t}^{1}\lambda_2^{2(i-1)}(W)MSE(t-i)\\
&= \lambda_2^{2t}\lim_{N\to\infty}\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}+O(f(\bar{\Delta}))+\frac{2\xi}{\Psi}\lambda_2^{2(t-1)}(W)u(t-1)*MSE(t)
\end{aligned} \qquad (D.23)$$

where $*$ denotes convolution and $u(t)$ is the discrete-time unit step function. We can simplify this expression further using frequency-domain techniques. Let $M(z) \triangleq \sum_{k=0}^{\infty} MSE(k)z^{-k}$. Then, taking the unilateral Z-transform of both sides of (D.23) and rearranging terms, we have that

$$M(z) \leq \frac{\lim_{N\to\infty} \frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}(1-z^{-1}) + O(f(\bar{\Delta}))(1-\lambda_2^2(W)z^{-1})}{(1-(\lambda_2^2(W)+\frac{2\xi}{\Psi})z^{-1})(1-z^{-1})}. \tag{D.24}$$

The inverse Z-transform of this expression yields the solution

$$\begin{aligned}\lim_{N\to\infty} MSE(t) \leq & \frac{O(f(\bar{\Delta}))(1-\lambda_2^2(W))}{1-\lambda_2^2(W)-\frac{2\xi}{\Psi}}u(t) \\ & + \left(\lim_{N\to\infty}\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} - \frac{\frac{2\xi}{\Psi}O(f(\bar{\Delta}))}{1-\lambda_2^2(W)-\frac{2\xi}{\Psi}}\right)\left(\lambda_2^2(W)+\frac{2\xi}{\Psi}\right)^t.\end{aligned} \tag{D.25}$$

For the convergence of this solution, the following must hold - (1) $\lambda_2^2(W)+\frac{2\xi}{\Psi} < 1$ and (2) $\lim_{N\to\infty}\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} - \frac{\frac{2\xi}{\Psi}O(f(\bar{\Delta}))}{1-\lambda_2^2(W)-\frac{2\xi}{\Psi}} \geq 0$. These conditions are equivalent to those stated in the lemma and they can be satisfied by choosing Ψ appropriately. This completes the proof. ■

D.2 Proof of Corollary 7.

Proof $\boldsymbol{\beta}(t) \triangleq \bar{\mathbf{x}}(t) - J\mathbf{x}(t)$. Using the recursion (D.3), $\boldsymbol{\beta}(t)$ can be written as

$$\boldsymbol{\beta}(t) = (W^t - J)\mathbf{x}(0) + \sum_{\ell=t}^{1}(W-J)^{\ell}\mathbf{v}(t-\ell) + \mathbf{v}(t) + \sum_{\ell=t}^{1}(W-J)^{\ell-1}\mathbf{e}(t-\ell).$$

Observe that upon taking the expectation of $\frac{1}{N}||\boldsymbol{\beta}(t)||^2$ we will get an expression which resembles equations (D.4)-(D.5) closely. So using similar arguments as before, one can derive the following recursion

$$\frac{\mathbb{E}\{||\boldsymbol{\beta}(t)||^2\}}{N} \leq \lambda_2^{2t}\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N}u(t) + bu(t) + \frac{2\xi'}{\Psi}\lambda_2^{2(t-1)}(W)u(t-1) * \frac{\mathbb{E}\{||\boldsymbol{\beta}(t)||^2\}}{N}$$

where $*$ denotes convolution, $u(t)$ denotes the discrete-time unit step function, and constant $b = O\left(\frac{1}{N}\right) + O\left(\frac{N_0}{\Psi N}\right) + O\left(\frac{N_0^2}{\Psi N^2}\right) + f(\bar{\Delta})\left(1 + \sum_{\ell=t}^{1} \lambda_2^{2\ell}(W)\right)$. We have used $\frac{\mathbb{E}\{||\mathbf{v}(t)||^2\}}{N} = \frac{\Delta^2}{12} = f(\bar{\Delta})$. Let $B(z) \triangleq \frac{1}{N}\sum_{k=0}^{\infty} \mathbb{E}\{||\boldsymbol{\beta}(k)||^2\} z^{-k}$. The recursion can be solved using frequency domain techniques to obtain the solution-

$$\begin{aligned} \frac{\mathbb{E}\{||\boldsymbol{\beta}(t)||^2\}}{N} &\leq \frac{b(1-\lambda_2^2(W))}{1-\lambda_2^2(W)-\frac{2\xi'}{\Psi}} u(t) \\ &\quad + \left(\frac{\mathbb{E}\{||\mathbf{x}(0)||^2\}}{N} - \frac{\frac{2\xi'}{\Psi} b}{1-\lambda_2^2(W)-\frac{2\xi'}{\Psi}}\right)\left(\lambda_2^2(W) + \frac{2\xi'}{\Psi}\right)^t. \end{aligned} \tag{D.26}$$

The solution converges under similar conditions as Lemma 21. ■

BIBLIOGRAPHY

[1] Jamshid Abouei, Alireza Bayesteh, and Amir K. Khandani. Delay-throughput analysis in decentralized single-hop wireless networks. In *Proc. IEEE Intl. Symp. Info. Th. (ISIT)*, Nice, France, June 2007.

[2] Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li, and Raymond Yeung. Network information flow. *IEEE Trans. Info. Theory*, 46(4):1204 – 1216, July 2000.

[3] O. Alay, P. Liu, Y. Wang, E. Erkip, and S. Panwar. Error resilient video multicast using randomized distributed space-time codes. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, March 2010.

[4] Animashree Anandkumar and L. Tong. Type-based random access for distributed detection over multiaccess fading channels. In *IEEE Transactions on Signal Processing*, volume 55, pages 5032 – 5043, October 2007.

[5] A. Salman Avestimehr, Suhas Diggavi, and David N. C. Tse. Wireless network information flow: A deterministic approach. submitted to IEEE Trans. Info. Th., 2010.

[6] T. C. Aysal, M. Coates, and M. Rabbat. Distributed average consensus using probabilistic quantization. *Proc. 14th IEEE Workshop on Statistical Signal Processing*, pages 640–644, August 2007.

[7] Marjan Baghaie and Bhaskar Krishnamachari. Fast flooding using cooperative transmissions in wireless networks. In *Proc. IEEE Intl. Conf. on Comm. (ICC)*, Dresden, Germany, June 2009.

[8] Nikhil Bansal and Zhen Liu. Capacity, delay and mobility in wireless ad-hoc networks. In *Proc. IEEE INFOCOM*, 2003.

[9] V. Blondel, J. Hendrickx, A. Olshevsky, and J. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. *Proc. of 44th IEEE Conference on Decision and Control*, 2005.

[10] Shashibhushan Borade, Lizhong Zheng, and Robert Gallager. Amplify-and-forward in wireless relay networks: Rate, diversity, and network size. *IEEE Trans. Info. Theory*, 53(10), Oct. 2007.

[11] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *Special Issue IEEE Trans. Information Theory and IEEE/ACM Transactions on Networking*, 52(6):2508 – 2530, June 2006.

[12] Viveck R. Cadambe and Syed Ali Jafar. Interference alignment and degrees of freedom of the k-user interference channel. *IEEE Trans. Info. Theory*, 54(8):3425 – 3441, August 2008.

[13] Ruggero Carli, Fabio Fagnani, Paolo Frasca, and Sandro Zampieri. Gossip consensus algorithms via quantized communication. *Automatica*, 2009.

[14] Ruggero Carli, Fabio Fagnani, Alberto Speranzon, and Sandro Zampieri. Communication constraints in the average consensus problem. *Automatica*, 44(3):671 – 684, 2008.

[15] Shiou-Hung Chen, Urbashi Mitra, and Bhaskar Krishnamachari. Cooperative communication and routing over fading channels in wireless sensor networks. In *IEEE International Conference on Wireless Networks,Communications, and Mobile Computing*, Maui, Hawaii, June 2005.

[16] Song Yean Cho, Cédric Adjih, and Philippe Jacquet. Heuristics for network coding in wireless networks. In *3rd Intl. Conf. on Wireless internet*, pages 1 – 5, 2007.

[17] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 1991.

[18] Tao Cui, Tracey Ho, and Jorg Kliewer. Memoryless relay strategies for two-way relay channels. *IEEE Transactions on Communications*, 57(10):3132 – 3143, Oct. 2009.

[19] G. Cybenko. Load balancing for distributed memory multiprocessors. *Journal of Parallel and Distributed Computing*, 7:279 – 301, 1989.

[20] Philip J. Davis. *Circulant Matrices*. New York: Wiley-Interscience, 1979.

[21] Alexandros G. Dimakis, Anand D. Sarwate, and Martin J. Wainwright. Geographic gossip: Efficient aggregation for sensor networks. In *Proc. IPSN*, Nashville, TN, April 2006.

[22] Fabio Fagnani and Sandro Zampieri. Randomized consensus algorithms over

large scale networks. *IEEE Journal on Selected Areas in Communications*, 6:634 – 649, 2008.

[23] Abbas El Gamal, James Mammen, Balaji Prabhakar, and Devavrat Shah. Throughput-delay trade-off in wireless networks. In *Proc. IEEE INFOCOM*, 2004.

[24] R. Gandhi, A. Mishra, and S. Parthasarathy. Minimizing broadcast latency and redundancy in ad hoc networks. *IEEE/ACM Trans. on Networking*, 16(4):840 – 851, August 2008.

[25] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46:388 – 404, March 2000.

[26] Thomas Halford and Keith Chugg. Barrage relay networks. In *Information Theory and Applications Workshop*, San Diego, CA, February 2010.

[27] Patrick Herhold, Emesto Zimmermann, and Gerhard Fettweis. A simple cooperative extension to wireless relaying. In *International Zurich Seminar On Communications*, Zurich, Switzerland, February 2004.

[28] Xiaoyan Hong, Kaixin Xu, and Mario Gerla. Scalable routing protocols for mobile ad hoc networks. *IEEE Network*, 16(4):11 – 21, July 2002.

[29] Y. W. Hong and A. Scaglione. On multiple access for correlated: a content-based group testing approach. In *Proc. IEEE Information Theory Workshop*, San Antonio, TX, October 2004.

[30] Yao Win Hong and Anna Scaglione. Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks. *IEEE Trans. on Wireless Communications*, 5(10):2844 – 2855, Oct. 2006.

[31] Scott Huang, Peng-Jun Wan, Jing Deng, and Yunghsiang S. Han. Broadcast scheduling in interference environment. *IEEE Trans. on Mobile Computing*, 7(11):1338 – 1348, 2008.

[32] A. Jain, S. Kulkarni, and S. Verdú. Multicasting in large random wireless networks: Bounds on the minimum energy per bit. In *IEEE Intl. Symposium on Information Theory*, pages 2627 – 2631, June 2009.

[33] Gentian Jakllari, Srikanth V. Krishnamurthy, Michalis Faloutsos, and Prashant V. Krishnamurthy. On broadcasting with cooperative diversity in

multi-hop wireless networks. *IEEE Journal on Selected Areas in Communications*, 25(2):484 – 496, February 2007.

[34] Sang-Woon Jeon and Sae-Young Chung. Two-phase opportunistic broadcasting in large wireless networks. In *Proc. IEEE Intl. Symp. Info. Th. (ISIT)*, pages 2771 – 2775, Nice, France, June 2007.

[35] Aravind Kailas and Mary Ingram. OLA with transmission threshold for strip networks. In *Proc. Military Communications Conference (MILCOM)*, Boston, MA, USA, Oct. 2009.

[36] Aman Kansal, Michel Goraczko, and Feng Zhao. Building a sensor network of mobile phones. In *Proc. of the 6th Intl. Conf. on Information processing in sensor networks*, pages 547 – 548, Cambridge, MA, 2007.

[37] S. Kar and Jose Moura. Distributed consensus algorithms in sensor networks: Quantized data and random link failures. *IEEE Trans. Signal Processing*, 2009.

[38] Soumya Kar and Jose Moura. Distributed average consensus in sensor networks with quantized inter-sensor communication. In *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, USA, April 2008.

[39] Akshay Kashyap, Tamer Başar, and R. Srikant. Quantized consensus. *Automatica*, 43(7):1192–1203, 2007.

[40] Sachin Katti, Shyamnath Gollakota, and Dina Katabi. Embracing wireless interference: analog network coding. In *SIGCOMM '07*, Kyoto, Japan, 2007.

[41] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proc. IEEE Symposium on Fundations of Computer Science*, page 482, 2003.

[42] A. Keshavarz-Haddad, V. Ribeiro, and Rudolf Reidi. Broadcast capacity in multihop wireless networks. In *Proc. ACM MOBIHOC*, pages 239 – 250, Los Angeles, CA, 2006.

[43] Zhenning Kong and Edmung Yeh. On the latency for information dissemination in mobile wireless networks. In *Proc. ACM MOBIHOC*, Hong Kong, May 2008.

[44] J. Korner and K. Marton. How to encode the modulo-two sum of binary sources. In *IEEE Trans. Info. Theory*, volume 25, pages 219 – 221, March 1979.

[45] J. Nicholas Laneman, David N. C. Tse, and Gregory W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Transactions on Information Theory*, 50:3062–3080, December 2004.

[46] K. Langendoen, A. Baggio, and O. Visser. Murphy loves potatoes: experiences from a pilot sensor network deployment in precision agriculture. In *Proc. Parallel and Distributed Processing Symposium*, April 2006.

[47] K. Liu and A.M. Sayeed. Type-based decentralized detection in wireless sensor networks. *IEEE Transactions on Signal Processing*, May 2007.

[48] I. Maric and R.D. Yates. Cooperative multihop broadcast for wireless networks. *IEEE Journal on Selected Areas in Communications*, 22:1080–1088, August 2004.

[49] Ivana Maric and Roy Yates. Cooperative multihop broadcast for wireless networks. *IEEE JSAC Special Issue on Fundamental Performance Limits of Wireless Sensor Networks*, 22(6), August 2004.

[50] Mortada Mehyar, Demetri Spanos, John Pongsajapan, Steven Low, and R.M. Murray. Asynchronous distributed averaging on communication networks. *to appear in IEEE Trans. on Networking*, August 2007.

[51] Birsen Sirkeci Mergen and Anna Scaglione. A continuum approach to dense wireless networks with cooperation. In *Proc. of INFOCOM 2005*, Miami, FL, March 2005.

[52] Birsen Sirkeci Mergen and Anna Scaglione. Randomized space-time coding for distributed cooperative communication. In *IEEE Transactions on Signal Processing*, submitted for publication.

[53] G. Mergen, V. Naware, and L. Tong. Asymptotic detection performance of type-based multiple access in sensor networks. In *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2005.

[54] G. Mergen and L. Tong. Type based estimation over multiaccess channels. *IEEE Trans. Signal Processing*, 54(2):613 – 626, February 2006.

[55] S. Muthukrishnan, B. Ghosh, and M. Schultz. First and second order diffusive methods for rapid, coarse, distributed load balancing. *Theory of Computing Systems*, 31:331 – 354, 1998.

[56] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath. Ergodic interference alignment. In *IEEE Intl. Symposium on Information Theory*, Seoul, Korea, June 2009.

[57] Bobak Nazer and Michael Gastpar. The case for structured random codes in network capacity theorems. In *European Transactions on Telecommunications, Special Issue on New Directions in Information Theory*, 2008.

[58] Michael J. Neely and Eytan Modiano. Capacity and delay tradeoffs for ad-hoc mobile networks. In *IEEE Trans. Info. Theory*, volume 51, June 2005.

[59] Sze-Yao Ni, Yu-Chee Tseng, Yuh-Shyan Chen, and Jang-Ping Sheu. The broadcast storm problem in a mobile ad hoc network. In *5th ACM/IEEE Intl. Conf. on Mobile computing and networking*, pages 151 – 162, Seattle, Washington, United States, 1999.

[60] R. Olfati-Saber. Consensus filters for sensor networks and distributed sensor fusion. In *Proc. 44th IEEE Conference on Decision and Control*, pages 6698 – 6703, Dec. 2005.

[61] R. Olfati-Saber. Distributed kalman filter with embedded consensus filters. In *Proc. 44th IEEE Conference on Decision and Control*, pages 8179 – 8184, Dec. 2005.

[62] R. Olfati-Saber and R. Murray. Agreement problems in networks with directed graphs and switching topology. In *Proc. 42nd Conference on Decision and Control*, 2003.

[63] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. on Automatic Control*, 49(9):1520 – 1533, Sept. 2004.

[64] Ayfer Özgür, Olivier Lévêque, and David N. C. Tse. Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks. *IEEE Trans. Info. Theory*, 53(10):3549 – 3572, Oct. 2007.

[65] Pulin Patel and Jack Holtzman. Analysis of a simple successive interference

cancellation scheme in a ds/cdma system. *IEEE Journal on Selected Areas in Communications*, 12(5):796 – 807, June 1994.

[66] B. Pittel. On spreading a rumor. *SIAM J. Appl. Math.*, 47:213–223, 1987.

[67] Petar Popovski and Hiroyuki Yomo. Physical network coding in two–way wireless relay channels. In *Proc. IEEE Intl. Conf. on Comm. (ICC)*, pages 707 – 712, 2007.

[68] J. G. Proakis. *Digital Comunications -4th ed.* McGraw-Hill, 2001.

[69] A. Qayyum, L. Viennot, and A. Laouiti. Multipoint relaying for flooding broadcast messages in mobile wireless networks. In *Proc. 35th Hawaii Intl. Conf.*, pages 3866 – 3875, January 2002.

[70] Wei Ren, Randal Beard, and Ella Atkins. A survey of consensus problems in multi-agent coordination. In *American Control Conference*, pages 1859 – 1864, Portland, OR, June 2005.

[71] Giovanni Resta and Paolo Santi. Latency and capacity optimal broadcasting in wireless multihop networks. In *Proc. IEEE Intl. Conf. on Comm. (ICC)*, Dresden, Germany, June 2009.

[72] Y. Sasson, D. Cavin, and A. Schiper. Probabilistic broadcast for flooding in wireless mobile ad hoc networks. In *Proc. Wireless Communications and Networking*, volume 2, pages 1124 – 1130, March 2003.

[73] Anna Scaglione and Yao-Win Hong. Opportunistic large arrays: Cooperative transmission in wireless multihop ad hoc networks to reach far distances. *IEEE transactions on Signal Processing*, 51:2082–2092, August 2003.

[74] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang. User cooperation diversity - part i: System description. *IEEE Trans. on Communications*, 21(11):1927–1938, November 2003.

[75] Masoud Sharif and Babak Hassibi. On the capacity of mimo broadcast channels with partial side information. *IEEE Trans. Information Theory*, 51(2), February 2005.

[76] Gaurav Sharma, Ravi Mazumdar, and Ness B. Shroff. Delay and capacity trade-offs in mobile ad hoc networks: A global perspective. In *IEEE/ACM Trans. on Networking*, volume 15, pages 981 – 992, Oct. 2007.

[77] B. Sirkeci-Mergen, A. Scaglione, and G. Mergen. Asymptotic analysis of multistage cooperative broadcast in wireless networks. *IEEE Trans. Information Theory*, 52(11):2531–2550, 2006.

[78] Ivana Stojanovic, Zeyu Wu, Masoud Sharif, and David Starobinski. Data dissemination in wireless broadcast channels: network coding versus cooperation. *IEEE Trans. on Wireless Communications*, 8(4):1726 – 1732, 2009.

[79] I. Stojmenovic, M. Seddigh, and J. Zunic. Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks. *IEEE Trans. Parallel and Distributed Systems*, 13(1):14 – 25, January 2001.

[80] L. Tong, Q. Zhao, and G. Mergen. Multipacket reception in random access wireless networks: from signal processing to optimal medium access control. *IEEE Communications Magazine*, 39(11):108 – 112, November 2001.

[81] S. Toumpis. Mother nature knows best: A survey of recent results on wireless networks based on analogies with physics. *Computer Networks*, 52(2):360–383, 2008.

[82] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[83] J. Tsitsikilis. *Problems in Decentralized Decision Making and Computation*. PhD thesis, MIT, 1984.

[84] S. Vanka, V. Gupta, and M. Haenggi. Power-delay analysis of consensus algorithms on wireless networks with interference. *Intl. Journal of Systems, Control, and Communications*, 2:256 – 274, 2010.

[85] S. Vanka, M. Haenggi, and V. Gupta. Distributed averaging in dense wireless networks. In *IEEE GLOBECOM*, Honolulu, HI, Dec. 2009.

[86] Mario Čagalj, Jean-Pierre Hubaux, and Christian Enz. Minimum-energy broadcast in all-wireless networks: Np-completeness and distribution issues. In *Proc. 8th Intl. Conf. on Mobile computing and networking (Mobicom)*, pages 172–182, Atlanta, GA, 2002.

[87] Pramod Viswanath, David N. C. Tse, and Rajiv Laroia. Opportunistic beamforming using dumb antennas. *IEEE Trans. Info. Theory*, 48(6):1277 – 1294, June 2002.

[88] Dilip Warrier and Upamanyu Madhow. On the capacity of cellular cdma with successive decoding and controlled power disparities. In *48th Vehicular Techonologies Conference (VTC)*, volume 3, pages 1873 – 1877, Ottawa, ON, Canada, 1998.

[89] S. Weber, J. Andrews, X. Yang, and G. de Veciana. Transmission capacity of cdma ad hoc networks employing successive interference cancellation. *IEEE Trans. Info. Theory*, 53(8), August 2007.

[90] Jeffrey Wieselthier, Gam D. Nguyen, and Anthony Ephremides. Energy-efficient broadcast and multicast trees in wireless networks. *Mobile Networks and Applications*, 7(6):481 – 492, 2002.

[91] Brad Williams and Tracy Camp. Comparison of broadcasting techniques for mobile ad hoc networks. In *MOBIHOC*, Laussane, Switzerland, June 2002.

[92] John M. Wozencraft and Irwin M. Jacobs. *Principles of Communciation Engineering*. John Wiley & Sons, Inc., 1965.

[93] L. Xiao, S. Boyd, and S.-J. Kim. Distributed average consensus with least-mean-square deviation. *Journal of Parallel and Distributed Computing*, 67(1):33 – 46, 2007.

[94] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Proc. International Conference on Information Processing in Sensor Networks*, Los Angeles, CA, April 2005.

[95] M. E. Yildiz and A. Scaglione. Differential nested lattice encoding for consensus problems. In *Proc. of IEEE Information Processing in Sensor Networks (IPSN)*, Boston, MA, USA, April 2007.

[96] M. E. Yildiz and A. Scaglione. Coding with side information for rate constrained consensus. *IEEE Trans. Signal Processing*, 2008.

[97] Shengli Zhang, Soung Chang Liew, and Patrick Lam. Physical-layer network coding. In *Proc. 12th Intl. Conf. on Mobile computing and networking*, pages 358 – 365, Los Angeles, CA, USA, 2006.